Spatial dependence in German labor markets

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Dipl.-Volksw. Franziska Lottmann

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Ulrich Kamecke

Gutachter:

1. Prof. Dr. Nikolaus Hautsch
2. Prof. Michael Burda, Ph.D.

Abgabedatum: 18. Januar 2013
Tag des Kolloquiums: 10. Juli 2013

Abstract

In this dissertation, we present different empirical analyses of regional labor markets in Germany. To account for the spatial structure of labor market activities, we apply spatial econometric methods to regional labor market data. In the first analysis, we propose a spatial panel model for German matching functions to avoid possibly biased and inefficient estimates due to spatial dependence. Based on an official data set, we show that neglecting spatial dependencies in the data results in upward-biased coefficients. Furthermore, our results suggest that a dynamic modeling is more appropriate for matching functions than a static approach. In the second analysis, we study determinants for regional differences in unemployment rates. We specify a spatial panel model to avoid biased and inefficient estimates due to spatial dependence. The study covers the whole of Germany as well as East and West Germany separately. Our results suggest that a spatial dynamic panel model is the best model for this analysis. Moreover, we find that German regional unemployment is of disequilibrium nature, which justifies political interventions. Finally, we study the spatial weights matrix which is the key component in spatial econometric models. We investigate empirically the issue of defining spatial weights in labor market applications and propose factors driving spatial dependence in regional labor markets. In addition to geographic distance, we consider different dimensions of economic distance as transmission channel of spatial dependence. To decide which factors influence spatial dependence in labor markets, we apply a higher-order spatial autoregressive model to data on regional labor markets in Germany. Our results suggest that geographic distance does not capture the spatial dependence between regional labor markets sufficiently but economic distance needs to be considered as well.

Zusammenfassung

Diese Dissertation umfasst verschiedene empirische Analysen regionaler Arbeitsmärkte in Deutschland. Wir wenden dabei Methoden der räumlichen Ökonometrie auf regionale Arbeitsmarktdaten an, um der räumlichen Struktur von Arbeitsmarktaktivitäten Rechnung zu tragen. In der ersten Analyse schlagen wir ein räumliches Paneldatenmodell zur Untersuchung deutscher Matchingfunktionen vor. Mit Hilfe dieses Modells sollen verzerrte und ineffiziente Koeffizientenschätzungen aufgrund von räumlichen Abhängigkeiten vermieden werden. Wir zeigen auf der Basis von amtlichen Daten, dass das Vernachlässigen der räumlichen Struktur zu nach oben verzerrten Matchingkoeffizienten führt. Zudem zeigen unsere Ergebnisse, dass ein dynamischer im Vergleich zu einem statischen Modellansatz für Matchingfunktionen besser geeignet ist. Das Ziel der zweiten Untersuchung ist es, Bestimmungsfaktoren für regionale Unterschiede in Arbeitslosigkeit zu identifizieren. Dafür spezifizieren wir ein räumliches Paneldatenmodell. Unsere Ergebnisse zeigen, dass ein räumlich und zeitlich dynamisches Paneldatenmodell am besten für diese Fragestellung geeignet ist. Weiterhin zeigen unsere Ergebnisse, dass die regionalen Unterschiede in der deutschen Arbeitslosigkeit einen Ungleichgewichtszustand darstellen. Diese Erkenntnis kann als Argument für politische Interventionen dienen. In der letzten Analyse wenden wir uns der räumlichen Gewichtungsmatrix zu, der eine zentrale Bedeutung in räumlichen Modellen zukommt. Auf Basis einer empirischen Analyse wollen wir die Definition von räumlichen Gewichten untersuchen. In diesem Zusammenhang ermitteln wir Faktoren, die die räumlichen Abhngigkeiten auf Arbeitsmärkten bestimmen. Dabei untersuchen wir sowohl unterschiedliche Dimensionen ökonomischer als auch geographische Distanzen als Wirkungskanal rumlicher Abhängigkeit. Für die Entscheidung, welche dieser Distanzdimensionen einen Einfluss auf die räumlichen Relationen hat, verwenden wir ein räumlich-autoregressives Modell höherer Ordnung. Unsere Ergebnisse zeigen, dass geographische Distanz alleine nicht ausreicht, um die räumlichen Interdependenzen zwischen regionalen Arbeitsmärkten zu erklären, sondern auch Dimensionen ökonomischer Distanz einen signifikanten Erklärgehalt haben.

Contents

List of Figures

List of Tables

Chapter 1

Spatial dependencies in German matching functions

This chapter is based on Lottmann (2012b).

1.1 Introduction

In 2009, there were about 9.25 million people that became unemployed in Germany. But, during the same time, about 9 million people left the state of inactivity while the average unemployment stock amounted to 3.42 million in 2009. These numbers illustrate that labor markets are characterized by large flows between the states of activity and inactivity. In macroeconomic research, a standard tool to analyze these dynamics is the matching function which describes how the flow of new hires (matches) is related to the unemployment stock and to the stock of vacancies. The matching function allows to analyze the determinants of job creation and the structure of underlying search frictions in labor markets.

However, as shown in this paper, labor market activity is correlated over space. The presence of spatial (auto-)correlation implies that the extent of matching in one particular region is correlated with that in neighboring regions. Neglecting spatial correlation when modeling the matching process yields biased and inefficient estimates of the matching function. This is widely ignored in the empirical matching literature as matching functions are mostly specified according to models assuming cross-sectional independence among observations. This independence assumption is questionable with respect to both the labor supply and the labor demand side. On the one hand, the search behavior of workers is not limited to one particular region resulting in migration and commuting of workers. On the other, the agglomeration literature shows that there are economies of scale due to spatial concentration of activity of firms within industries (see, for example, Ciccone and Hall (1996)). According to Rosenthal and Strange (2001), one reason for agglomeration is labor pooling, i.e. firms of the same industry tend to cluster in space in order to profit from a pool of specialized workers in this region.

The aim of this paper is the estimation of matching functions taking into account spatial dependencies in order to obtain unbiased and efficient estimates. For the estimation, we use an official data set that provides monthly information on 176 local employment offices (*Arbeitsagenturen*) for the period from 2000 until 2009. To exploit the panel structure of the data, we specify the matching function using a spatial panel model. In addition to a static model, we use a dynamic model for the matching function to capture the positive (temporal) autocorrelation in the data. Most contributions in the empirical matching literature apply only a static modeling to the matching function. The combination of both spatial econometric methods and dynamic modeling is novel to this literature.

The estimation of matching functions has been subject of intensive research in the literature. In their seminal paper, Blanchard and Diamond (1989) estimate matching functions using aggregated time series for the United States. After that, other authors provide studies on aggregated matching functions for different countries. Van Ours (1991) analyzes the Netherlands, Berman (1997) estimates aggregate matching functions for Israel and Burda and Wyplosz (1994) investigate the labor markets of Germany, Spain, France and the United Kingdom. Approaches that utilize aggregated time series assume that the national economy acts as a single labor market (Coles and Smith (1996)). Due to many factors that hamper mobility, as individual preferences, social ties, differences in real income, etc., it is more reasonable to consider the national economy as a collection of spatially distinct labor markets (Coles and Smith (1996)). Therefore, authors turned to estimating matching functions using regional data sets. Burda (1993) uses data on Czech and Slovak employment offices, Coles and Smith (1996) use regional labor market data for the United Kingdom and Anderson and Burgess (2000) use state-level data for the United States. However, these contributions do not model cross-sectional dependencies explicitly. The contributions by Fahr and Sunde (2001, 2005, 2006a, 2006b, 2009) also deal with data on German labor markets. Our paper extends the range of their analysis by using data for both West and East Germany covering a more recent period.

To our best knowledge, only a few contributions deal with spatial dependencies in the empirical matching context as Burgess and Profit (2001), Hynninen (2005), Fahr and Sunde (2006a, b) and Dmitrijeva (2008). These authors introduce spatial interactions into their model using spatially lagged exogenous variables. This is a simple way of modeling a spatial process since estimation of such models can be done using standard estimation techniques. As suggested by test results on cross-sectional dependence in the residuals, this model does not capture the spatial autocorrelation in the data in a sufficient way. Therefore, we apply panel models including a spatial lag and a spatial error term to the matching function. Lee and Yu (2010c) propose a quasi-maximum likelihood approach for the static spatial autoregressive panel data model with fixed effects which we adopt here. For the estimation of the dynamic model, we employ the estimation methodology suggested by Lee and Yu (2010a).

An important component of spatial econometric modeling is the spatial weights matrix. In addition to the binary contiguity matrix that is mostly used in the literature, we exploit a data set on commuting relations between local employment offices to construct both binary spatial weights matrices with entries zero and one and spatial weights matrices with general weights. We argue that the amount of commuting captures the actual spatial relations on labor markets better than the binary contiguity matrix.

We extend the existing literature by the following three aspects: Firstly, we estimate matching functions controlling for both temporal and spatial (auto-)correlation by applying recent estimation methodologies for spatial panel models. Secondly, we find that ignoring spatial dependencies in matching data when modeling the matching function results in upward-biased matching elasticities. As the estimated matching elasticities reflect the structural features of the matching process, this finding is important. Thirdly, our results suggest that compared to a static model, a dynamic approach results in a better fit of the data.

The structure of the paper is as follows: The second section presents the basic matching model while the third presents the data set and explains how the spatial weights matrix is defined. In order to motivate the spatial econometric approach, the fourth section provides test results of the (global) Moran I test for spatial autocorrelation. Section five presents the econometric model and the sixth section is dedicated to the estimation results. Finally, the last section concludes.

1.2 Matching on labor markets

In macroeconomics, the matching function plays a central role for the analysis of labor market dynamics and labor market efficiency. Petrongolo and Pissarides (2001) state that "it occupies the same place in the macroeconomist's tool kit as other aggregate functions, such as the production function and the demand for money function". The labor market is assumed to be a decentralized market where it takes time and resources for unemployed individuals and vacant jobs to find each other. Reasons for this complicated exchange process are trading frictions, incomplete information and heterogeneities. With the help of the matching function, this two-sided search process can be characterized. Petrongolo and Pissarides (2001) provide a survey on the empirical matching function literature in which they discuss different matching function specifications and the results.

In the empirical matching literature, it is standard to use a Cobb-Douglas specification for the matching function.[1] Taking logs, the matching equation describing the flow

[1]From a theoretical point of view, it is also possible to use a CES-type matching function. In this context, Burda (1994) explains that the assumption of this type of matching function does not entail additional explanatory power. Nevertheless, there are critical views concerning the Cobb-Douglas assumption for matching functions in the literature, see, for example, Stevens (2007).

of matches m_{it} between time period t and $t+1$ is given by

$$ln\, m_{it} = c_i + \alpha_t + \beta_1\, ln\, \mathcal{U}_{it} + \beta_2\, ln\, \mathcal{V}_{it} + \epsilon_{it}, \quad t = 1, \ldots, T, i = 1, \ldots, n \tag{1.1}$$

where $\mathcal{U}_{it}$ and $\mathcal{V}_{it}$ denote the stock of registered unemployment and the stock of registered vacancies at time point t, respectively.[2] c_i is the time-invariant effect controlling for employment office-specific characteristics as, for example, its size, while α_t is a time effect controlling for aggregate shocks. ϵ_{it} describes the error term which is assumed to be homoscedastic and uncorrelated.[3]

The basic matching model (equation 1.1) ignores any spatial dependence effects. If spatial dependencies in the data are neglected, standard OLS regression will provide biased parameter estimates in case of spatial lag dependence and in case of spatially lagged exogenous variables. Though, OLS regression produces unbiased and inefficient estimates for the spatial error model. Neglecting the spatial lag term is similar to an omitted variables bias (Franzese and Hays (2007)). As the spatial lag term is always correlated with the errors, OLS estimation of the corresponding coefficient will be inconsistent (see Anselin and Bera (1998) or Franzese and Hays (2007)).

1.3 Data and spatial weights matrix

1.3.1 Measuring matches, unemployment and vacancies

We use monthly data on unemployment, vacancies and matches for the period from 2000 until 2009. This data is provided by the Federal Employment Office (*Bundesagentur für Arbeit*) for all local employment offices in Germany. The allocation of local employment offices is done by the Federal Employment Office according to administrative reasons. Except for some changes in the assignment of local employment offices in Berlin,[4] the allocation of local employment offices is stable over time. Unemployed are registered and coached by the local employment office at their place of residence. The 176 local employment offices constitute an exhaustive sample of Germany as they cover the whole country.

We use the "outflows from unemployment into gainful employment" as measure for the matches. The "outflows from unemployment" in general also include people

[2] In order to ease notation for the spatial panel models, we distinguish between stocks and flows by using this notation in script for the stocks.

[3] Note that this matching function specification follows the random matching approach, i.e. agents are matched randomly at any point in time, independent of their duration of search. Contrary to this, the stock-flow approach assumes that unemployed individuals have complete information about possible jobs. Either they find a job instantaneously or they wait until new vacancies arrive on the market. Therefore, the stock-flow approach does not only take into account stocks but also inflows into unemployment and vacancies.

[4] During the period from 2000 to 2009, some of Berlin's local employment offices have been merged. In order to have the same structure throughout our study period, all local employment offices of Berlin have been merged to one.

entering into part-time employment, into labor-market policy measures or people leaving the labor force which we do not want to consider as successful matches. Firms are not obliged to report their vacant jobs to the Federal Employment Office in Germany. Therefore, the registered vacancies represent only a fraction of the overall economic supply of vacant positions. In 2006, this fraction amounted to 44% only (BA (2008)). The unemployment data is collected in accordance to the "concept of registered unemployment" which is regulated in the German Social Security Code. Hence, this analysis is limited to that part of the labor market which is officially registered at the Federal Employment Office. However, registered positions can also be filled with employed job searchers which are not covered in our data set. The registered unemployed and vacancies are possibly subject to a downward skill bias. On the one hand, highly-qualified individuals mostly do not use the Federal Employment Office in order to find a new job. On the other hand, firms having vacant jobs for which a high qualification is needed prefer using web portals, national newspapers and internal channels to find suitable candidates (Koppel (2008)). Likewise, Christensen (2001) argues that the rate of reported vacancies is higher for jobs which require low skills.

Table 1.1 shows the summary statistics of our panel data set over time. The unemployment stock increases steadily between 2000 and 2005 while it is decreasing afterwards.[5] Finally, the unemployment stock increases slightly in 2009. The summary statistics show strong regional variation in the unemployment stocks with values for the standard deviation of more than 20,000. Note that this data does not control for the size of the local employment offices. The local employment office of Berlin is bigger by construction which explains the high maximum values. The matching data does not show a clear trend and its variation is the lowest compared to the stock variables. Finally, the number of vacancies decreases until 2004 while it is increasing between 2004 and 2007. It decreases again during the last two years of our sample period. Note that this data does not necessarily reflect the evolution of the overall labor demand but it reflects when and how many jobs are reported by firms.

We divide the matches in every local employment office by corresponding unemployment stocks for the computation of the test statistics for spatial autocorrelation. Hence, the resulting matching-unemployment ratios represent the fraction of jobless individuals leaving unemployment in order to start a job.

1.3.2 Time series properties

To test for stationarity of the data, we apply panel unit root tests. The results of the Im et al. (2003) test and the Fisher-type (ADF) test, that was proposed by Maddala and Wu (1999) as well as Choi (2001), clearly reject the hypothesis of a unit root in the unemployment and vacancy data as all p-values are zero. For the matches, the hypothesis of

[5]The peak of unemployed individuals in 2005 can be explained by a labor market reform ("*Hartz* IV reform") which changed the definition of who is considered as unemployed. See section 1.6 for more details on this reform.

a unit root can only be rejected in case of the Fisher-type test. A more detailed description of these tests can be found in Appendix A. Nevertheless, it has to be noted that the results of Baltagi et al. (2007a) show that there can be considerable size distortions in panel unit root tests when the true model exhibits spatial error correlation. Therefore, the test results can only serve as an indication of possible nonstationarities in the data.

To analyze the time series properties of our data, we produce autocorrelation function (ACF) plots for the 176 time series of the three variables. As most of them exhibit a similar correlation structure, we only show some representative examples here (Figures 1.1 - 1.6). Figures 1.1 and 1.2 show the two different correlation schemes of the matches where in one case the correlation is long-lasting (more than two years) and in the other it dissipates after one year. The correlation in unemployment stocks is either steadily decreasing, as shown in Figure 1.4, or it is long-lasting and simultaneously slightly double U-shaped (Figure 1.3). Finally, the correlation in vacancy stocks either dissipates quickly (after about 5 months) or it is U-shaped during one year with negative correlation afterwards (Figures 1.5 and 1.6).

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Unemployment stock:										
Min	3186	3318	3921	4991	4975	5309	5005	3630	2643	2970
1st quartile	10290	10210	11250	12370	12380	13620	12700	10220	8419	9358
Median	15910	15900	17300	18390	18420	20780	19620	16050	13880	14550
Mean	22130	21940	23030	24760	24830	27280	25780	21690	18690	19360
3rd quartile	28250	28240	29270	31230	31170	33870	32550	27550	23970	23920
Max	328200	304000	298500	331500	328900	332900	313600	278900	252300	245300
standard deviation	24320.46	23808.45	23675.53	24892.48	25364.8	27482.16	25707.97	22825.81	20424.71	20251.61
Matches:										
Min	222	211	233	235	268	303	250	257	255	313
1st quartile	795	744	770,8	856	819	849,8	839,8	808	770	814
Median	1170	1118	1119	1224	1164	1202	1224	1166	1130	1174
Mean	1487	1448	1436	1556	1475	1569	1576	1506	1467	1523
3rd quartile	1761	1692	1687	1847	1729	1774	1815	1770	1724	1760
Max	13750	19660	17120	14730	17550	20680	19680	18800	20320	20030
standard deviation	1221.85	1396.09	1250.50	1241.40	1321.22	1491.49	1443.39	1371.79	1446.74	1490.66
Vacancy stocks:										
Min	260	292	259	185	146	146	153	214	296	218
1st quartile	1491	1487	1369	1019	793	1014	1433	1606	1503	1267
Median	2254	2168	1952	1445	1152	1535	2162	2378	2226	1830
Mean	2956	2880	2590	1998	1651	2175	2900	3279	3091	2728
3rd quartile	3343	3285	2938	2298	1859	2548	3350	3750	3512	3100
Max	34600	36840	36170	38140	37030	38980	39000	38630	41440	36720
standard deviation	3036.23	2886.85	2541.75	2023.35	1883.03	2479.48	3159.17	3332.89	3243.17	3199.69

Source: Federal Employment Office

Table 1.1 Yearly summary statistics of unemployment stock, matches and vacancy stock of German local employment offices (2000-2009)

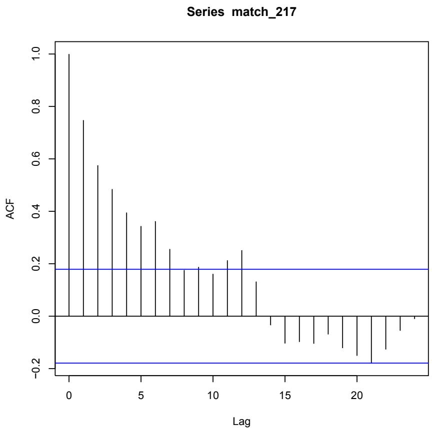


Figure 1.1 ACF plot of the matches in Bremen with a maximum lag length of 24 months (2000-2009)

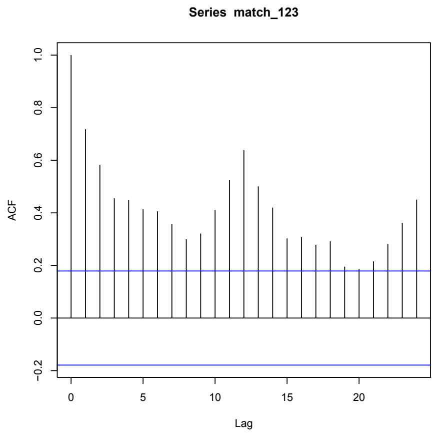


Figure 1.2 ACF plot of the matches in Hamburg with a maximum lag length of 24 months (2000-2009)

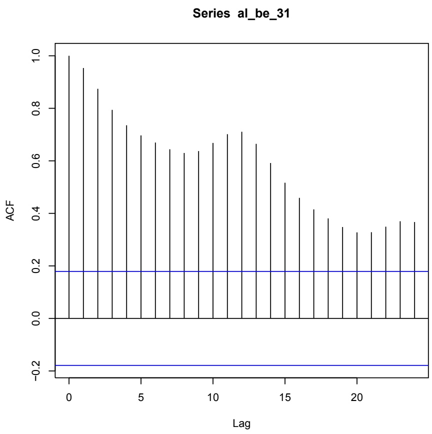


Figure 1.3 ACF plot of the unemployment stock in Mecklenburg-Western Pomerania with a maximum lag length of 24 months (2000-2009)

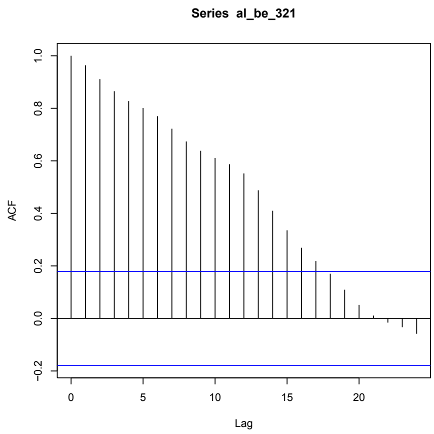


Figure 1.4 ACF plot of the unemployment stock in North-Rhine Westphalia with a maximum lag length of 24 months (2000-2009)

1.3.3 Specification of spatial influence

A fundamental building block of spatial econometric modeling is the spatial weights matrix. It is a nonstochastic matrix which defines exogenously the neighborhood of a certain location. Hence, the term 'neighboring' in the present context addresses the neighborhood set which is defined by the corresponding spatial weights matrix. We

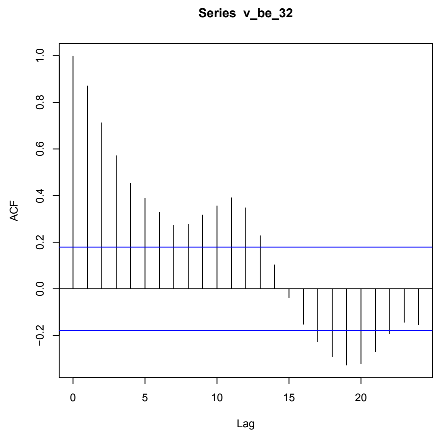


Figure 1.5 ACF plot of the vacancy stock in Mecklenburg-Western Pomerania with a maximum lag length of 24 months (2000-2009)

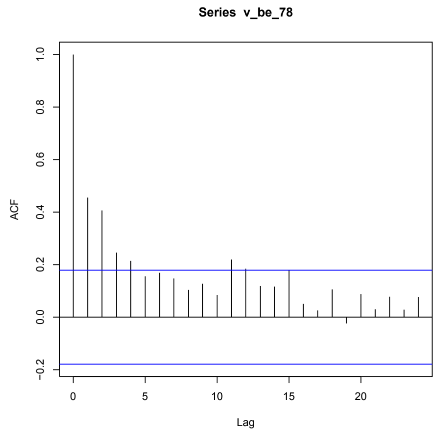


Figure 1.6 ACF plot of the vacancy stock in Saxony with a maximum lag length of 24 months (2000-2009)

use both binary spatial weights matrices where the entries are either zero or one and matrices with general weights.

The simplest version of a binary spatial weights matrix is a contiguity matrix which is mostly used in the empirical matching literature. When two local employment offices are neighbors, i.e. they share a common border, the corresponding entry in the matrix is one and zero otherwise. The elements on the main diagonal are zero by construction. This matrix induces a simple spatial structure which might be not sufficient to capture the actual spatial relations on German labor markets. Commuting of people is not limited to the neighboring region and, additionally, the binary contiguity matrix weights all neighbors equally. The latter assumption might be critical for a region that is surrounded by both a big city and a rural area. In this case, one would suspect that more people commute to the big city than to the rural area. A distance decay function, as used by Burgess and Profit (2001), is also not able to reproduce this pattern of spatial relations on labor markets because two regions can be separated by the same distance while having a different impact (in labor market sense) on each other.

To address these problems, we exploit a data set on commuting relations between different local employment offices. The amount of commuting reflects differences in labor market opportunities between local employment offices. Moreover, it measures the mobility of employees. Möller and Aldashev (2007), who also use commuter streams for constructing a spatial weights matrix, state that such a matrix captures the strength of interregional relationships among labor markets. The commuting data is also collected by the Federal Employment Office and it is part of the official statistic as well. It records all people who have a job that is subject to social insurance contribution. The numbers of commuters between local employment offices are recorded yearly at the appointed date June 30th. Our data set covers the period from 2000 to 2009. We use this data as a proxy for the interregional linkages between local employment offices. Therefore, we construct the average commuter matrix $P = (p_{ij})$, $i, j = 1, \ldots, 176$. The element p_{ij} of this matrix indicates the average number of people that live in employment office j and work in employment office i. Hence, row i of the average commuter matrix P contains the incoming commuters to region i while the elements of column j represent the outgoing commuters from employment office j to all other regions. The summary statistics of matrix P are found in Table 1.2. They show that there is commuting between most of the local employment offices, although it is not very strong between 75% of them. The highest numbers of incoming commuters are attained in big cities as Berlin, Düsseldorf, Cologne, Frankfurt, Munich and Hamburg.

Min	1st quartile	Median	Mean	3rd. quartile	Max	Std. dev.
0	6.2	16.3	216.02	45.8	60201.5	1583.24

Source: Federal Employment Office, author's calculations

Table 1.2 Summary statistics of the average commuter matrix P (2000-2009)

We use the commuting information twofold to construct both binary weights matrices and weights matrices with general weights. First, we discretize the information and construct additional binary spatial weights matrices. To control for the size of local employment offices, we divide the entries of the average commuter matrix by the average working age population.[6] We consider two local employment offices as neighbors when the commuting flow from region j to i divided by the working age population in region i exceeds a certain value δ. Hence, the entries of the spatial weights matrix are defined by

$$w_{ij} = \begin{cases} 0, & p_{ij}/pop_i^{15-64} < \delta \\ 1, & p_{ij}/pop_i^{15-64} \geq \delta, \end{cases} \tag{1.2}$$

where pop_i^{15-64} denotes the average working age population in region i. Regarding the choice of the cutoff value δ, we experimented with different values. However, we only present results of those values for which the resulting spatial weights matrix contains at least one neighbor for each region. The resulting spatial weights matrix still weights all neighbors equally, yet it reflects the actual connections on the labor market in a better way by not restricting the analysis to physical neighbors.

Second, we exploit the full information contained in the average commuter matrix P to construct a spatial weights matrix with general weights. Contrary to the applied spatial econometric literature where a distance decay function is often assumed, we need a monotonically increasing function as more intense commuting implies stronger spatial influence. We use the linear function for the specification of the weights, i.e.

[6] In official statistics of Germany, the working age is defined from 15 till 64 years.

$w_{ij} = p_{ij}$.[7] This function implies that the marginal influence of one additional commuter is constant.[8]

When computing spatially lagged matches, i.e. $Wln(M_{it})$, with the help of the general spatial weights matrix, the neighboring matches are weighted by the rows of the spatial weights matrix, i.e. by incoming commuters to region i. In general, this weighting scheme can be changed so that the neighboring matches are weighted by the outgoing commuters of region i. We also implemented this weighting scheme in our regressions and got virtually the same results.

1.4 Spatial dependencies in German labor markets

1.4.1 Empirical evidence on (global) spatial autocorrelation

A standard test for spatial autocorrelation is the Moran I test which was developed by Moran (1950). This test is not specified for a particular spatial process. Its null hypothesis is the absence of spatial autocorrelation whereas the alternative is not exactly specified. The test statistic can be expressed by

$$I = \frac{n}{S_0} \frac{e'We}{e'e}, \tag{1.3}$$

where $e = y - X\tilde{\beta}$ is a vector of standard OLS regression residuals, $\tilde{\beta} = (X'X)^{-1}X'y$, W denotes the spatial weights matrix and n is the number of observations (Anselin and Bera (1998)). In our case, y are the matches and the matrix X contains unemployment and vacancy stocks. S_0 is a standardization factor which is equal to the sum of the spatial weights, i.e. $S_0 = \sum_i \sum_j w_{ij}$. Cliff and Ord (1981) show that I is asymptotically normally distributed for normally distributed regression residuals. Therefore, inference is based on the standard normal variate $z(I)$ which is yielded by the transformation $z(I) = \frac{I - E(I)}{\sqrt{V(I)}}$. The expectation $E(I)$ and the variance $V(I)$ are derived by Cliff and Ord (1972) under the null hypothesis of no spatial dependence.

Since Moran's I test is designed to detect spatial autocorrelation from cross-section residuals, the test statistic is computed using matching-unemployment ratios for each

[7] We also considered the quadratic and the logarithmic function to construct the weights. The results of the weights produced by the quadratic function are always worse in the sense of Akaike's and the Bayesian information criterion and they are therefore not presented for the sake of brevity. However, they can be obtained from the author upon request. The logarithmic function produces results that give an indication for the spatial process to be nonstationary. This finding is supported by high values of the global Moran I statistic which, according to Fingleton (1999), can serve as an indicator for spatial nonstationarity.

[8] In principle, one could also divide the entries of the general spatial weights matrix by the working age population. But as we need to standardize the spatial weights matrix such that the row sums equal unity for our estimation procedure, this yields the same standardized matrix.

month within the period from 2000 to 2009. The values of the global Moran *I* statistic are positive and significant on all reasonable significance levels for all months within the period. Hence, we conclude that the regional distribution of matching-unemployment ratios in Germany is characterized by strong spatial dependencies. Figure 1.7 shows the evolution of the Moran *I* values for matching-unemployment ratios, its nine-month moving average and a linear trend line using the binary spatial weights matrix over the period from 2000 to 2009.

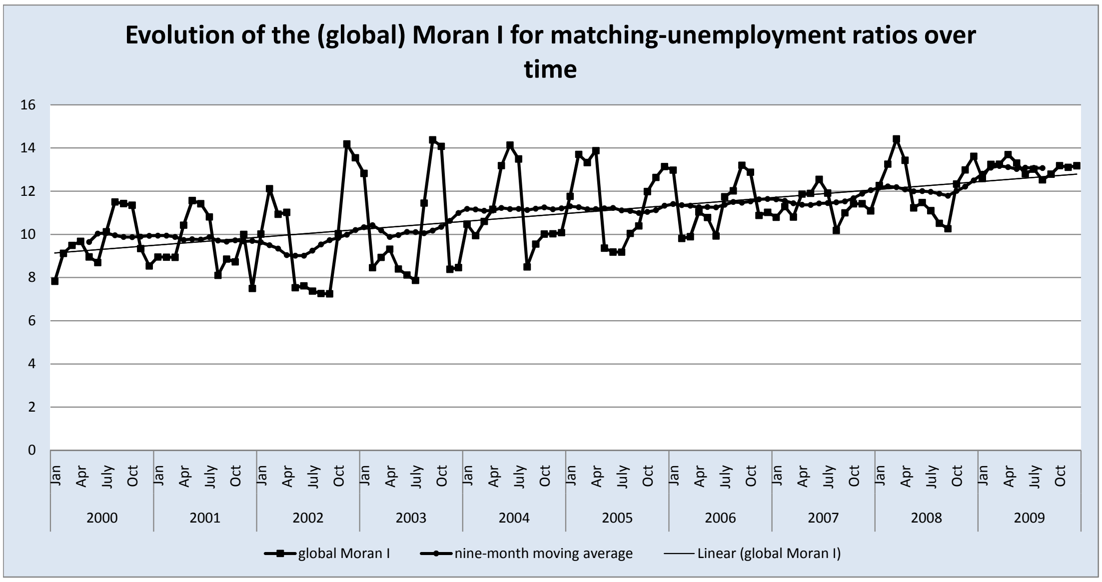
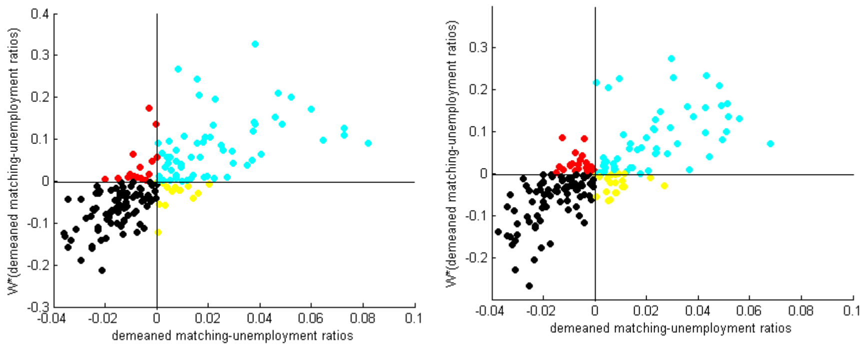


Figure 1.7 Evolution of the global Moran *I* for standardized matches for the period from 2000 until 2009

The linear trend is positive which means that spatial integration of German local employment offices becomes stronger during the period from 2000 until 2009. A reason for this is the increased mobility of people which is also supported by our commuting data. The mean relative change of (incoming) commuters between 2000 and 2009 amounts to 0.17, i.e. commuting increases significantly during this period.

Furthermore, Figure 1.7 shows some seasonal pattern in the (global) Moran values. The spatial autocorrelation is the strongest in April and October. As the matches itself also attain the highest values during spring time, we can conclude that spatial dependencies seem to be stronger when the labor market is more active. Additionally, the strong spatial autocorrelation in October shows that there are unobserved factors in the matching process that are correlated over space. Such factors can be business cycle



Note: This figure plots spatially lagged demeaned matching-unemployment ratios against demeaned matching-unemployment ratios. The colors identify the location of the points in the four quadrants. They are used to show where the points are located on the German map (see Figure 1.9).

Figure 1.8 Moran scatter plot using yearly averages of matching-unemployment ratios in 2000 and 2009

effects, the search behavior of unemployed and firms and differences in qualification.[9]

1.4.2 Local structure of spatial autocorrelation

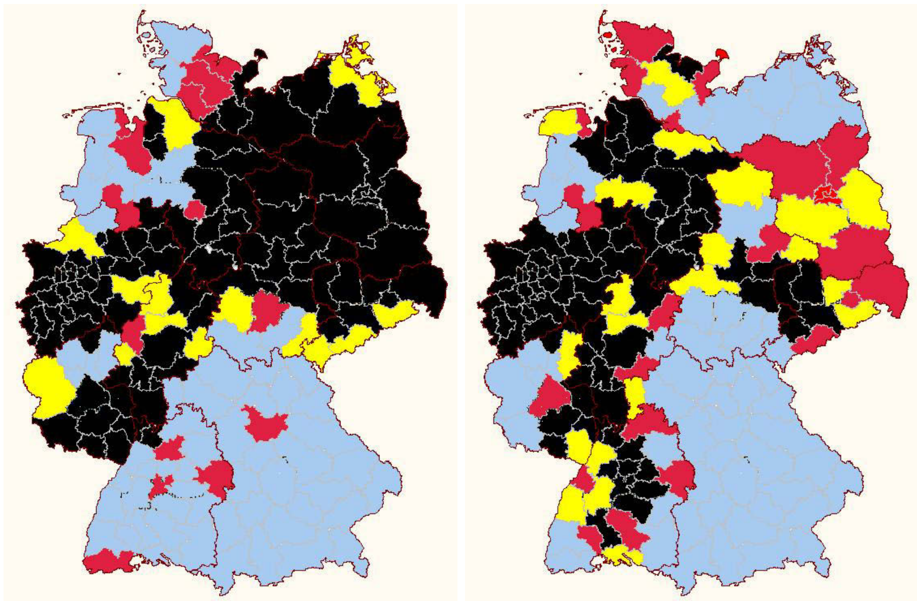
The (global) Moran I test only gives information about the global pattern of spatial dependence holding for all local employment offices in Germany. In order to analyze the local pattern of spatial autocorrelation, we compute Moran scatter plots (Anselin et al. (1996b)). These are based on the interpretation of the Moran I statistic as a regression coefficient in a regression of Wy on y where y denotes demeaned matching-unemployment ratios in the present analysis. In order to show this, Anselin et al. (1996b) rearranges the Moran I statistic to get the expression

$$I = \frac{y'Wy}{y'y} \tag{1.4}$$

which holds for a row-standardized spatial weights matrix, i.e. $S_0 = n$. Using the interpretation of the (global) Moran I statistic as a regression coefficient, the linear relationship between y and Wy can be visualized by a bivariate scatter plot of Wy against y.

The Moran scatter plots for the matching-unemployment ratios (yearly averages) of the years 2000 and 2009 are shown in Figure 1.8. They show that most of the local employment offices are positively spatially autocorrelated since most of the points lie

[9] Burgess and Profit (2001) analyze the cyclical variation of spatial dependence in the matching context using British data and find that the intensity of spatial dependence for unemployment outflows moves counter-cyclically. Their explanation is that unemployed individuals lower their search radius while firms have to search more widely in good times.



Note: The colors indicate the four quadrants of the Moran scatter plot as shown in Figure 1.8.

Figure 1.9 German map indicating the position of points in Moran scatter plot for yearly averages of matching-unemployment ratios (2000 and 2009)

in the first and third quadrant. This is in line with the results of the global Moran *I* test. The position of local employment offices in the first and third quadrant indicates that local employment offices with above-mean matching-unemployment ratios have neighbors with the same characteristic, while local employment offices with below-mean matching-unemployment ratios are more likely to be surrounded by local employment offices with low values. The remaining points in the second and fourth quadrant represent local employment offices which are negatively spatially autocorrelated.

Figure 1.9 shows maps of Germany indicating the location of the points in the Moran scatter plot in 2000 and 2009 (Figure 1.8). Interestingly, the Moran maps do not clearly replicate the former border between East and West Germany. In 2000, they show a band of local employment offices from Western (North-Rhine Westphalia) to Eastern Germany (Brandenburg) which seems to be positively spatially autocorrelated with matching-unemployment ratios below the mean. Most of the south German local employment offices are positively spatially autocorrelated with above-mean matching-unemployment ratios. Contrary to this, the situation diversifies in 2009. In the northern part of East Germany, the matching-unemployment ratios are positively spatially autocorrelated with above-mean values while there are local employment offices in Brandenburg and Saxony that are negatively spatially autocorrelated. In the south-west part of Germany, the matching-unemployment ratios changed from mostly being positively spatially autocorrelated with above-mean values to being negatively spatially autocorrelated as well as to positively spatially autocorrelated values with below-mean values.

Finally, the Moran scatter plots also support the fact that German matching data exhibit spatial autocorrelation. Thus, we have to take into account this fact in our econometric analysis.

1.5 Econometric Modeling

To capture the spatial dependence and the panel structure of the data, we propose to model the matching function by a spatial panel model. Since we do not have a representative sample of German employment offices but data on all local employment offices in Germany, a fixed effects model is preferred. In order to control for aggregate shocks, we use a model that takes into account time effects. Following most contributions in the empirical matching literature, we use a static specification of the matching function. Beyond that, we also specify the matching function in a dynamic way to capture the (temporal) autocorrelation of the data.

1.5.1 Static model specification

Our static model specification contains a spatial lag of the dependent variable as well as a spatial process for the error term. The corresponding matching equation is given by

$$\begin{aligned} ln\, M_t &= \lambda\, W\, ln\, M_t + \beta_1\, ln\, \mathcal{U}_t + \beta_2\, ln\, \mathcal{V}_t + c_n + \alpha_t \mathbb{1}_n + \Omega_t, \\ \Omega_t &= \rho W \Omega_t + \Xi_t, t = 1, \ldots, T, \end{aligned} \tag{1.5}$$

where $M_t = (m_{1t}, m_{2t}, \ldots, m_{nt})'$ is the $(n \times 1)$ vector of matches, $\mathcal{U}_t$ and $\mathcal{V}_t$ are the $(n \times 1)$ vectors of the unemployment and vacancy stocks, respectively. c_n represents the $(n \times 1)$ vector of fixed individual effects and α_t is the fixed time effect. W is the $(n \times n)$ nonstochastic spatial weights matrix and $\mathbb{1}_n$ is a $(n \times 1)$ vector of ones. $\Xi_t = (\xi_{1t}, \xi_{2t}, \ldots, \xi_{nt})'$ represents the $(n \times 1)$ vector of errors for which it is assumed that ξ_{it} are i.i.d. across i and t with zero mean and constant variance σ^2.

A spatial error term implies that there are spatially correlated random components influencing a region of more than one local employment offices. Examples in the labor market context are regional shocks as changes in regional governments or the closure of a production site. The spatial lag structure implies that the matching process in a particular local employment office is influenced by matching in other locations. As matching theory suggests, the matches are determined by the unemployment and vacancy stock. Thus, the spatial influence of all variables is captured implicitly by using the spatial lag model.

The way of inserting spatial autocorrelation into the matching function goes beyond most matching specifications in the literature that control for spatial dependencies because we use a spatial lag and spatial error term in our static model. Contrary to this, in the empirical matching literature, spatial dependencies are incorporated by spatially lagged exogenous variables into matching functions. In this way, the external effect of unemployment and vacancies on the matching process in neighboring local employment offices can be estimated. As these additional regressors are exogenous and as the error term remains spherical, estimation by ordinary least squares is unbiased and consistent (see Klotz (2004) for the pooled case). We also adopted this modeling to our data and got insignificant spatial spillovers of both stock variables. Moreover, we apply Pesaran's CD test (Pesaran (2004)) to test for cross-sectional dependence in the residuals. The results show that there is spatial correlation left in the residuals, i.e. the model with spatially lagged exogenous variables incompletely captures the spatial autocorrelation in the data.[10]

Lee and Yu (2010c) propose a quasi-maximum likelihood approach for the estimation of model (1.5). They show that (direct) maximum likelihood estimation yields inconsistent parameter estimates (unless n is large). Even in the case when n and T are large, the asymptotic distribution of the estimates is not properly centered. Therefore, they propose a transformation approach to eliminate the individual and time effects. The transformations are the deviation from time mean, $J_T = I_T - \frac{1}{T}\mathbb{1}_T\mathbb{1}_T'$, and the deviation from cross section mean, $J_n = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n'$, operator as used in the literature on panel data analysis (see Baltagi (2005)). The disturbance terms in the resulting equation (after performing these operations) would be linearly dependent. For this reason, their proposition is to base the transformations on the orthonormal eigenvector matrices of J_T and J_n. Let $[F_{T,T-1}, \frac{1}{\sqrt{T}}\mathbb{1}_T]$ be the orthonormal eigenvector matrix of J_T where $F_{T,T-1}$ is the $(T \times (T-1))$ submatrix corresponding to eigenvalues of one. Furthermore, let $[F_{n,n-1}, \frac{1}{\sqrt{n}}\mathbb{1}_n]$ be the orthonormal eigenvector matrix of J_n where $F_{n,n-1}$ is the $(n \times (n-1))$ submatrix corresponding to eigenvalues of one. The matching function (1.5) is firstly transformed by $F_{T,T-1}$ which yields

$$\begin{aligned}(ln\, M_t)^* &= \lambda W(ln\, M_t)^* + \beta_1(ln\, \mathcal{U}_t)^* + \beta_2(ln\, \mathcal{V}_t)^* + \alpha_t^*\mathbb{1}_n + \Omega_t^* \\ \Omega_t^* &= \rho W\Omega_t^* + \Xi_t^*, t = 1, \ldots, T-1,\end{aligned} \tag{1.6}$$

where $(ln\, M_t)^* = [ln\, M_{n1}, \ldots, ln\, M_{nT}]F_{T,T-1}$ and $[\alpha_1^*\mathbb{1}_n, \alpha_2^*\mathbb{1}_n, \ldots, \alpha_{T-1}^*\mathbb{1}_n] = [\alpha_1\mathbb{1}_n, \alpha_2\mathbb{1}_n, \ldots, \alpha_T\mathbb{1}_n]'F_{T,T-1}$ are transformed time effects. Secondly, in order to eliminate the time effects, the model is further transformed by $F_{n,n-1}$ yielding a $(n-1)$-dimensional vector $(ln\, M_t)^{**}$ such that $(ln\, M_t)^{**} = F_{n,n-1}'(ln\, M_t)^*$, i.e.

$$\begin{aligned}(ln\, M_t)^{**} &= \lambda(F_{n,n-1}'WF_{n,n-1})(ln\, M_t)^{**} + \beta_1(ln\, \mathcal{U}_t)^{**} + \beta_2(ln\, \mathcal{V}_t)^{**} + \Omega_t^{**} \\ \Omega_t^{**} &= \rho(F_{n,n-1}'WF_{n,n-1})\Omega_t^{**} + \Xi_t^{**}, t = 1, \ldots, T-1,\end{aligned} \tag{1.7}$$

where $(ln\, \mathcal{U}_t)^{**} = F_{n,n-1}'(ln\, \mathcal{U}_t)^*$ and $(ln\, \mathcal{V}_t)^{**} = F_{n,n-1}'(ln\, \mathcal{V}_t)^*$. Note that the effective sample size after both transformations is $(n-1)(T-1)$ and that the spatial weights matrix needs to be row-normalized for this transformation approach.

[10]The full results of the estimation and the test can be obtained from the author upon request.

The transformed equation (1.7) can be estimated by quasi-maximum likelihood. After some rearrangements, Lee and Yu (2010c) derive the log-likelihood function for the transformed model (1.7)

$$\begin{aligned} ln\, L_{n,T}(\theta) = -\frac{(n-1)(T-1)}{2} ln\, 2\pi\sigma^2 - (T-1)[ln(1-\lambda) + ln(1-\rho)] \\ + (T-1)[ln|S_n(\lambda)| + ln|R_n(\rho)|] - \frac{1}{2\sigma^2}\sum_{t=1}^{T} \widetilde{\Xi}'_t(\phi) J_n \widetilde{\Xi}_t(\phi), \quad (1.8) \end{aligned}$$

where $\theta = (\beta', \lambda, \rho, \sigma^2)$, $\beta' = (\beta_1, \beta_2)'$, $\phi = (\beta', \lambda, \rho)'$, $S_n(\lambda) = I_n - \lambda W$, $R_n(\rho) = I_n - \rho W$ and $\widetilde{\Xi}_t = R_n(\rho)[S_n(\lambda)\widetilde{ln\, M_t} - (\widetilde{ln\, \mathcal{U}_t}, \widetilde{ln\, \mathcal{V}_t})\beta]$. Note that $\widetilde{ln\, M_t} = ln\, M_t - \overline{ln\, M_t}$ for $t = 1, \ldots, T$, where $\overline{ln\, M_t} = \frac{1}{T}\sum_{t=1}^{T} ln\, M_t$. $\widetilde{ln\, \mathcal{U}_t}$, $\widetilde{ln\, \mathcal{V}_t}$ and $\widetilde{\Xi}_t$ are defined analogously.

Lee and Yu (2010c) show that the resulting quasi maximum-likelihood estimates for all parameters are consistent when either $n \to \infty$ or $T \to \infty$ and asymptotically normally distributed. Additionally, they derive explicitly the asymptotic distribution and show that it is properly centered.

1.5.2 Dynamic model specification

As shown in the data section, labor market data exhibit positive temporal autocorrelation. To capture these dynamics, we apply a spatial dynamic panel data model. In addition to a temporally lagged term, it contains a spatial lag term and a combined spatially and temporally lagged term of the dependent variable. Applying this model to our matching function, yields

$$\begin{aligned} ln\, M_t = \lambda\, W\, ln\, M_t + \gamma\, ln\, M_{t-1} + \delta\, W\, ln\, M_{t-1} + \beta_1\, ln\, \mathcal{U}_t + \beta_2\, ln\, \mathcal{V}_t \\ + c_n + \alpha_t \mathbb{1}_n + \Xi_t, t = 1, \ldots, T, \quad (1.9) \end{aligned}$$

where γ captures the pure time-dynamic effect and δ captures the combined spatial-time effect. The assumptions about the error term Ξ_t are as before.

We adopt the methodology proposed in Lee and Yu (2010b) and Lee and Yu (2010a) for the estimation of model (1.9). Lee and Yu (2010b) show that (direct) maximum likelihood estimation yields a bias of order $O(max(1/n, 1/T))$ for the common parameters. Therefore, they propose two variants of a transformation approach to avoid this bias. On the one hand, the transformation J_n in combination with an eigenvalue and eigenvector decomposition is applied and, on the other, the model is transformed by $(I_n - W)$. Lee and Yu (2010b) show that the quasi-maximum likelihood estimates from the maximization of the log-likelihood function of the J_n-transformed model are free of $O(1/n)$ bias. Nevertheless, the resulting quasi-maximum likelihood estimates are biased and, therefore, Lee and Yu (2010a) propose a bias correction procedure which we also apply.

The $(I_n - W)$-transformation eliminates not only time effects but also possible unstable components. Thus, it can be applied to all possible data generating processes. We applied both transformations to our data. But as the results are fairly similar and in order to save space, we present only the results and theoretical foundations of the $(I_n - W)$-transformation approach.

Transforming the dynamic matching equation (1.9) by $(I_n - W)$, yields

$$(I_n - W)ln\, M_t = \lambda W(I_n - W)ln\, M_t + \gamma(I_n - W)ln\, M_{t-1} + \delta W(I_n - W)ln\, M_{t-1} \\ + (I_n - W)\mathcal{X}_t\beta + (I_n - W)c_n + (I_n - W)\Xi_t,\ t = 1, \ldots, T, \quad (1.10)$$

where $\mathcal{X}_t = [ln\, \mathcal{U}_t, ln\, \mathcal{V}_t]$. The variance-covariance matrix of $(I_n - W)\Xi_t$ is given by

$$Var((I_n - W)\Xi_t) = \sigma^2\Sigma_n \quad (1.11)$$

with $\Sigma_n = (I_n - W)(I_n - W)'$. As the components of the error term in the transformed model (1.10) are linearly dependent, an eigenvalue-eigenvector decomposition is used again. For that, the matrix $[F_n, H_n]$ is defined to be the orthonormal matrix of eigenvectors and Λ_n is defined to be the diagonal matrix of nonzero eigenvalues of Σ_n such that $\Sigma_n F_n = F_n\Lambda_n$ and $\Sigma_n H_n = 0$. The columns of F_n consist of eigenvectors corresponding to nonzero eigenvalues and those of H_n are for zero eigenvalues of Σ_n. According to Lee and Yu (2010a), the transformed spatial weights matrix is defined as $W^* = \Lambda_n^{-1/2}F_n'WF_n\Lambda_n^{-1/2}$. Then, the (further) transformed model is given by

$$(ln\, M_t)^* = \lambda W^*(ln\, M_t)^* + \gamma(ln\, M_{t-1})^* + \delta W^*(ln\, M_{t-1})^* + \mathcal{X}_t^*\beta + c_n^* + \Xi_t^*, \\ t = 1, \ldots, T, \quad (1.12)$$

where $(ln\, M_t)^* = \Lambda_n^{-1/2}F_n'(I_n - W)ln\, M_t$ and the other variables are defined accordingly. Note that the transformed vector $(ln\, M_t)^*$ is of dimension n^* where n^* is the rank of $\sigma^2\Sigma_n$. The concentrated log-likelihood of equation (1.12) is

$$ln\, L_{n,T}(\theta) = -\frac{n^*T}{2}ln\, 2\pi - \frac{n^*T}{2}ln\, \sigma^2 - (n - n^*)Tln(1 - \lambda) + Tln|S_n(\lambda)| \\ - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\widetilde{\Xi}_t'(\theta)(I_n - W)'\Sigma_n^+(I_n - W)\widetilde{\Xi}_t(\theta) \quad (1.13)$$

where $\widetilde{\Xi}_t(\theta) = S_n(\lambda)\widetilde{ln\, M_t} - \widetilde{Z}_t\vartheta$ with $Z_t = (ln\, M_{t-1}, Wln\, M_{t-1}, \mathcal{X}_t)$ and $\vartheta = (\gamma, \delta, \beta')$.

1.6 Empirical Results

In order to improve the success of the Federal Employment Office in placing jobless individuals in a job, the German government passed different laws to reform the German labor market during the period from 2002 until 2005 ("*Hartz* reforms"). These laws constitute the ”largest labor market reform in Germany in the post-war period in terms of speeding up the matching process between unemployed and vacant jobs” (Fahr and Sunde (2009)). The aim was to accelerate labor market flows and to reduce unemployment duration (Fahr and Sunde (2009)). Since one part of these reforms (becoming effective in 2005) entailed changes in the official definition of unemployment, meaning that more people are coached by the Federal Employment Office, we analyze the periods from 2000 until 2004 and from 2005 until 2009 separately.

Firstly, we estimate the basic matching model without any spatial terms. It is specified according to a two-way fixed effects model and it is estimated using the standard within-estimator.[11] Secondly, we estimate the static matching specification and, thirdly, the dynamic matching model, both using the different spatial weights matrices that we defined in Section 1.3.3. As we have five different specifications for the spatial weights matrix, we have eleven regressions for each period. The regression results are shown in Tables 1.3 and 1.5 for the period from 2000 until 2004 while the results for the second period are presented in Tables 1.4 and 1.6.

As suggested by matching theory, the estimated elasticities of matches on both stocks are positive and significant in all specifications. The elasticity of matches on unemployment for the basic model during the period from 2000 until 2004 amounts to 0.599. This means that an increase of the unemployment stock by 1% results in an increase of matching by 0.599 percent. The estimated elasticities with respect to vacancies are much smaller than those with respect to the unemployed in all specifications. This finding might be related to the underreporting of vacant jobs to the Federal Employment Office. Furthermore, it can be explained by the high vacancy turnover, i.e. vacant jobs are filled within a month and, thus, are not counted in the end-of-month stocks.

Ignoring spatial effects yields biased estimates of the elasticities on both stock variables. Our results show that the elasticities in the basic matching model are upward-biased. The existence of this bias is theoretically shown in Franzese and Hays (2007). They argue that neglecting a spatial lag process results in an omitted-variable bias. We show the consequences of this bias using some figures implied by the results. We compare the elasticity on the unemployment stock of the basic model with that of the best (according to information criteria) static model. During the period from 2000 until 2004, the average unemployment stock amounted to 23,335.06. A one per cent increase in unemployment corresponds to about 233 additional individuals. The average number of matches was 1481 and an increase by 0.599 % is equivalent to about nine additional matches. Thus, one new match of a vacant job and an unemployed is created when the unemployment stock increases by $233/9 \approx 26$ individuals. In case of the static model, using the binary spatial weights matrix that is constructed with respect to $\delta = 0.0005$, one new match is created when the number of unemployed increases by about 39 individuals. This simple calculation shows that the relative importance of job searchers in the matching process is overestimated when cross-sectional dependencies are not accounted for.

In the empirical matching literature, Fahr and Sunde (2006a, b) also use German

[11] For more details on this subject, see, for example, Baltagi (2005).

dependent variable: $ln\ M_{it}$ time period: 2000-2004						
	basic	static				
		contiguity	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\,\mathcal{U}_{it}$	0.599***	0.333***	0.307***	0.771***	0.382***	0.362***
	(30.9)	(16.8)	(15.25)	(40.18)	(19.61)	(18.41)
$ln\,\mathcal{V}_{it}$	0.092***	0.062***	0.066***	0.082***	0.067***	0.067***
	(16.4)	(11.32)	(11.81)	(15.22)	(12.87)	(12.39)
λ	—	0.035***	0.027***	0.068***	0.126***	0.051***
		(3.71)	(2.98)	(6.08)	(11.19)	(4.52)
ρ	—	0.054***	0.045***	0.091***	0.152***	0.07***
		(5.4)	(4.67)	(7.95)	(12.75)	(5.9)
σ^2	0.021	0.019***	0.02***	0.018***	0.017***	0.019***
		(84.64)	(726.04)	(105.46)	(141.75)	(82.59)
log-like	5624.658	5737.146	5557.677	5899.043	6281.426	5815.381
BIC	-0.857	-1.085	-1.063	-1.128	-1.201	-1.099
observations	10560	10560	10440	10440	10440	10560

Notes: t-statistics in parentheses. t-statistics of the static spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010c). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. The different sample sizes stem from the fact that the population data we used for the spatial weights matrices with a cutoff value is only available for 174 local employment offices. The local employment offices of the Saarland have been merged to one. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.3 Estimates of matching functions using the basic and the static spatial panel model for the period from 2000 until 2004

data. However, it is difficult to compare our results with their results as they use yearly data covering only Western Germany for the period from 1980 until 1997 in both contributions. Due to German reunification in 1990, the German economy grew by one forth bringing along substantial structural differences within the unified country. In addition to that, they use a different data source for the matching data.

The spatial autoregressive (λ) and the spatial autocorrelation (ρ) coefficient measuring the spatial effects in our model are significant and positive. Hence, the number of matches in the neighborhood influences the matching process in a particular local employment office. The positive spatial autocorrelation coefficient indicates regional effects that affect the matching process in more than one local employment office. The effect of the spatial error term is stronger than that of the spatial lag term for the static model, i.e. spatially correlated random components play an important role on German labor markets.

Furthermore, the estimation results show that the matching elasticities are fairly robust with respect to the choice of the spatial weights matrix which is more pronounced in case of the static model with one exception (period from 2000 until 2004, binary spatial weights matrix with cutoff value $\delta = 0.001$). Notably, this holds for the vacancies in all specifications. Likewise, the time-dynamic effect (γ) is not sensitive to different spatial regimes as well. However, this is not true for the estimates of the spatial coefficients

dependent variable: $ln\ M_{it}$						
time period: 2005-2009						
	basic	static				
		contiguity	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\,\mathcal{U}_{it}$	0.698***	0.519***	0.526***	0.563***	0.554***	0.533***
	(39.97)	(29.05)	(29.34)	(31.96)	(31.52)	(29.97)
$ln\,\mathcal{V}_{it}$	0.1***	0.074***	0.073***	0.079***	0.078***	0.077***
	(18.09)	(13.49)	(13.3)	(14.64)	(14.68)	(14.06)
λ	—	0.032***	0.028***	0.066***	0.123***	0.047***
		(2.8)	(2.68)	(5.24)	(8.79)	(3.57)
ρ	—	0.047***	0.044***	0.085***	0.144***	0.063***
		(3.93)	(3.96)	(6.44)	(9.69)	(4.52)
σ^2	0.019	0.019***	0.019***	0.018***	0.017***	0.018***
		(108.67)	(121.73)	(133.35)	(145.88)	(108.08)
log-like	5936.595	5933.831	5852.389	6076.891	6328.234	5997.738
BIC	-0.916	-1.122	-1.119	-1.162	-1.21	-1.134
observations	10560	10560	10440	10440	10440	10560

Notes: t-statistics are in parentheses. t-statistics of the static spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010c). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. The different sample sizes stem from the fact that the population data we used for the spatial weights matrices with a cutoff value is only available for 174 local employment offices. The local employment offices of the Saarland have been merged to one. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.4 Estimates of matching functions using the basic and the static spatial panel model for the period from 2005 until 2009

dependent variable: $ln\ M_{it}$ period: 2000-2004					
	dynamic (I-W)-transformation (after bias correction)				
	contiguity	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\,\mathcal{U}_{it}$	0.285^{***}	0.246^{***}	0.366^{***}	0.391^{***}	0.382^{***}
	(22.13)	(20.11)	(22.99)	(24.88)	(24.19)
$ln\,\mathcal{V}_{it}$	0.047^{***}	0.042^{***}	0.049^{***}	0.053^{***}	0.05^{***}
	(12.9)	(11.4)	(11.21)	(12.31)	(11.37)
λ	0.638^{***}	0.626^{***}	0.096^{***}	0.157^{***}	0.071^{***}
	(72.65)	(70.8)	(4.99)	(7.31)	(3.82)
γ	0.466^{***}	0.472^{***}	0.502^{***}	0.508^{***}	0.508^{***}
	(53.63)	(53.77)	(48.8)	(50.24)	(49.31)
δ	-0.231^{***}	-0.257^{***}	0.066^{***}	0.019	0.071^{***}
	(-17.81)	(-20.56)	(3.39)	(0.88)	(3.63)
σ^2	0.009^{***}	0.009^{***}	0.012^{***}	0.012^{***}	0.013^{***}
	(70.92)	(70.19)	(50.26)	(50.87)	(52.56)
log-like	9618.98	9077.575	8057.838	8229.887	8049.474
BIC	-1.82	-1.766	-1.567	-1.601	-1.522
observations	10559	10266	10266	10266	10559

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010a). λ is the spatial autoregressive coefficient, γ captures the time dynamic effect and δ captures the combined spatial-time effect. The different sample sizes stem from the fact that the population data we used for the spatial weights matrices with a cutoff value is only available for 174 local employment offices. The local employment offices of the Saarland have been merged to one. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.5 Estimates of matching functions using the spatial dynamic panel model for the period from 2000 until 2004

dependent variable: $ln\ M_{it}$ period: 2005-2009					
	dynamic (I-W)-transformation (after bias correction)				
	contiguity	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\,\mathcal{U}_{it}$	0.365^{***}	0.359^{***}	0.404^{***}	0.414^{***}	0.37^{***}
	(27.96)	(27.18)	(27.42)	(28.24)	(28.2)
$ln\ \mathcal{V}_{it}$	0.039^{***}	0.038^{***}	0.038^{***}	0.039^{***}	0.039^{***}
	(9.62)	(9.27)	(8.25)	(8.6)	(9.6)
λ	0.538^{***}	0.537^{***}	0.087^{***}	0.148^{***}	0.621^{***}
	(53.54)	(54.54)	(4.84)	(7.2)	(52.53)
γ	0.507^{***}	0.512^{***}	0.511^{***}	0.515^{***}	0.512^{***}
	(61.14)	(61.55)	(54.93)	(56)	(61.98)
δ	-0.246^{***}	-0.266^{***}	0.052^{***}	0.008	-0.316^{***}
	(-18.2)	(-20.88)	(2.75)	(0.38)	(-20.43)
σ^2	0.01^{***}	0.01^{***}	0.012^{***}	0.012^{***}	0.01^{***}
	(71.16)	(70.56)	(57.95)	(58.45)	(71.32)
log-like	9148.268	8907.953	8189.919	8319.399	9114.678
BIC	-1.731	-1.733	-1.593	-1.618	-1.724
observations	10559	10266	10266	10266	10559

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010a). λ is the spatial autoregressive coefficient, γ captures the time dynamic effect and δ captures the combined spatial-time effect. The different sample sizes stem from the fact that the population data we used for the spatial weights matrices with a cutoff value is only available for 174 local employment offices. The local employment offices of the Saarland have been merged to one. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.6 Estimates of matching functions using the spatial dynamic panel model for the period from 2005 until 2009

(λ and ρ) because they are sensitive to the choice of the spatial weights matrix.[12]

In the empirical matching literature, matching functions are mostly specified in a static way.[13] However, according to the Bayesian information criterion, the dynamic model fits the data better than the static model. This is in line with the positive temporal autocorrelation detected in the data. Thus, a dynamic approach is more appropriate for modeling matching functions.

Compared with the static model, the matching elasticities of the dynamic model are smaller. This can be explained by the strong time-dynamic effect in the data which is absorbed by the coefficients of the static model. Moreover, the coefficient of the spatial

[12]Hujer et al. (2009) also find that the long-term effect of labor market policies is unaffected by changes in the spatial weights while the estimates of spatial coefficients differ with the choice of the spatial weights matrix.

[13]One exception is the contribution by Hujer et al. (2009), for example, that specifies a spatial dynamic matching function for the analysis of the indirect and direct effects of active labor market policy at the regional level for Western Germany.

lag term (λ) is larger in the dynamic model. A reason for this might be that the dynamic model only contains a spatial lag term but not a spatial error term as the static model. The space-time effect is negative in some of the specifications. This means that an increase in the number of matches in neighboring local employment offices during the previous period results in lower matches during the present period. However, this result has to be interpreted carefully. According to Ochsen (2009), the negative sign can arise from the perfect correlation of the space-time lagged variable with the time lagged and the spatially lagged variables.

Comparing both subperiods, the estimated elasticities of matches with respect to unemployment are larger during the period from 2005 until 2009 which holds for both the static and the dynamic model. Hence, the effect of additional unemployment on matching is stronger. However, the picture for the elasticities with respect to vacancies is different: They are larger during the time from 2000 until 2004 in case of the dynamic model while the opposite holds for the static model. Hence, we can conclude for the dynamic model that the relation between the unemployment stock and vacant jobs has improved during the second period. The estimated spatial effects are similar for both subperiods in case of the static model while this is not true for the dynamic model. Only the pure time-dynamic effect is fairly similar in both subperiods.

Do bigger labor markets exhibit stronger matching?

Our matching function specifications allow for fixed effects, i.e. for employment office-specific characteristics as, for example, the size. However, the fixed effects specification does not control for the impact of regional agglomeration. Therefore, we want to construct an agglomeration index that captures information about the surrounding area of local employment office i. We expect the matching process to be more efficient in larger labor markets. For the construction of the agglomeration index, we again use our data on working age population. It shows strong variation between local employment offices. The smallest local employment office has a working age population of 63,540 people while in the biggest 2,436,000 people of working age are living. The median is 272,300 and the standard deviation amounts to about 238,579.[14] Using this data, we construct the agglomeration index proposed in Longhi et al. (2006) which is defined by

$$A_{it} = \sum_j pop_{jt}^{15-64} w_{ij} \tag{1.14}$$

where w_{ij} are the entries of the binary contiguity matrix. Hence, A_{it} is an indicator for the size of the labor market in the neighborhood of local employment office i. However, data on working age population is only available on a yearly basis. In order to compare results, Tables 1.7 - 1.10 show not only the estimation results with the agglomeration index but also without. Note that the time series dimension is heavily reduced in this estimation as this analysis is based on yearly information. We include the natural logarithm

[14]These numbers are again affected by the local employment office of Berlin which is bigger by construction.

	basic	static				
		binary	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\, \mathcal{U}_{it}$	0.013***	0.006	0.006	0.007	0.007	0.006*
	(3.11)	(1.61)	(1.76)	(1.52)	(1.56)	(1.65)
$ln\, \mathcal{V}_{it}$	0.051***	0.038***	0.042***	0.082***	0.079***	0.035***
	(5.51)	(3.11)	(3.38)	(9.33)	(9.1)	(4.75)
λ	—	0.767***	0.709***	-0.245	0.011	0.912***
		(5.07)	(4.16)	(-0.68)	(0.02)	(6.12)
ρ	—	-0.572	-0.418	0.762***	0.772***	-0.677
		(-1.61)	(-1.26)	(4.98)	(3.67)	(-1.53)
σ^2	0.009	0.006***	0.006***	0.006***	0.007***	0.006***
		(16.69)	(29.21)	(8.55)	(8.19)	(22.45)
log-like	1587.185	1614.614	1622.721	1660.202	1662.853	1633.659
BIC	-1.815	-1.847	-1.856	-1.899	-1.902	-1.868
obs.	1740	1740	1740	1740	1740	1740

dependent variable: $ln\ M_{it}$; time period: 2000-2009

Notes: t-statistics are in parentheses. t-statistics of the static spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010c). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.7 Estimates of matching functions for the basic and the static spatial panel model using yearly averages (2000-2009)

of A_{it} as additional explanatory variable in our regressions but we also experimented with A_{it} divided by 10^6 as in the original paper by Longhi et al. (2006) and interactions of A_{it} with explanatory variables, although these variables are not significant. We also considered the spatial interaction indicator, $T_{it} = \sum_j (pop_{jt}^{15-64} pop_{it}^{15-64})^{0.5} w_{ij}$, which is also proposed in Longhi et al. (2006), but it is also not significant.

The empirical results (Tables 1.8 and 1.10) show a significant and positive elasticity of the agglomeration index only in some of the static model specifications while it is not significant in the basic and the spatial dynamic model. The matching elasticities on both stock variables are not affected by the inclusion of the agglomeration index both in the static and the dynamic model while this is more pronounced in the dynamic model. In our yearly regressions, the unemployment stock is not significant in the spatial models. It seems that the spatial lag and the spatial error term incorporate the influence of the unemployment stock. The results of the static model give the impression that the agglomeration index captures part of the spatial autocorrelation since the coefficients measuring the spatial influence are smaller in most of the specifications and significant in a fewer number of cases. However, the results of the dynamic model do not support this because the estimated coefficients of the spatial parameters do not differ when including the agglomeration index. Regarding the information criterion, there is no clear difference between the models with and without the agglomeration index. Both the static and the dynamic model with the binary weights matrix constructed with respect to the cutoff value $\delta = 0.0005$ fit the data best.

	dependent variable: $ln\ M_{it}$					
	time period: 2000-2009					
	basic	static				
		binary	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\, \mathcal{U}_{it}$	0.013^{***}	0.006	0.004	0.007	0.007	0.003
	(3.1)	(1.39)	(1.17)	(0.6)	(0.98)	(1.09)
$ln\, \mathcal{V}_{it}$	0.051^{***}	0.086^{***}	0.055^{***}	0.083^{**}	0.079^{***}	0.048^{***}
	(5.51)	(9)	(3.77)	(2.17)	(5.29)	(3.81)
$ln\ A_{it}$	-0.001	0.113	0.545^{**}	0.683	0.269	0.43^{**}
	(-0.15)	(0.58)	(2.21)	(0.26)	(0.07)	(2.26)
λ	—	0.028	0.662^{***}	0.182	0.24	0.858^{***}
		(0.43)	(3.52)	(0.03)	(0.06)	(6.04)
ρ	—	0.041	-0.385	0.497	0.632	-0.619^{**}
		(0.52)	(-1.16)	(0.11)	(0.2)	(-2.08)
σ^2	0.009	0.008^{***}	0.006^{***}	0.007^{**}	0.007^{***}	0.006^{***}
		(6.08)	(26.85)	(2.37)	(5.14)	(9.74)
log-like	1586.196	1522.264	1639.087	1658.347	1662.319	1646.157
BIC	-1.816	-1.739	-1.873	-1.895	-1.9	-1.881
obs.	1740	1740	1740	1740	1740	1740

Notes: t-statistics are in parentheses. t-statistics of the static spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010c). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.8 Estimates of matching functions including the agglomeration index using the basic and the static spatial panel model (2000-2009)

dependent variable: $ln\ M_{it}$ period: 2000-2009					
	dynamic (I-W)-transformation (after bias correction)				
	binary	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\,\mathcal{U}_{it}$	0.003	0.003	0.002	0.002	0.003
	(1.02)	(1.02)	(0.81)	(0.64)	(1.03)
$ln\,\mathcal{V}_{it}$	0.028***	0.027***	0.03***	0.032***	0.027***
	(4.22)	(4.2)	(4.71)	(4.88)	(4.09)
λ	0.42***	0.429***	0.582***	0.701***	0.513***
	(14.26)	(15.59)	(17.05)	(17.31)	(15.2)
γ	0.583***	0.585***	0.57***	0.565***	0.582***
	(29.77)	(29.82)	(28.71)	(28.5)	(29.61)
δ	−0.199***	−0.213***	−0.269***	−0.301***	−0.238***
	(-5.28)	(-5.97)	(-6.07)	(-5.73)	(-5.51)
σ^2	0.004***	0.004***	0.004***	0.004***	0.004***
	(29.19)	(28.88)	(28.96)	(29.05)	(28.92)
log-like	2323.32	2339.535	2354.772	2357.995	2336.474
BIC	-2.955	-2.976	-2.995	-2.999	-2.972
obs.	1566	1566	1566	1566	1566

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010a). λ is the spatial autoregressive coefficient, γ captures the time dynamic effect and δ captures the combined spatial-time effect. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.9 Estimates of matching functions for the spatial dynamic panel model using yearly averages (2000-2009)

Controlling for border effects

As Germany is a member of the European Union (EU), the freedom of movement, one of the four fundamental freedoms of the EU, also applies to the German labor market. European Union citizens have the right to work and live in another member state without a special permit.[15] Hence, the local employment offices at the German border may be influenced by neighboring countries. Our commuting data set only contains information on people residing abroad and working in Germany. During the period from 2000 until 2009, on average, about 113,509 foreign people worked in Germany which is about 0.4 % of employment that is subject to social insurance contribution. But as commuting in the other direction might be more important in the matching context, we

[15]However, due to transitional arrangements, member states have the right to restrict access to their labor market for citizens of the new member states that acceded the European Union on May 1st, 2004 (Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovenia and Slovakia) and on 1 January 2007 (Bulgaria and Romania). Germany claimed this opportunity for most of the countries for the maximal period of seven years after accession of the corresponding country.

dependent variable: *ln* M_{it} period: 2000-2009					
	dynamic (I-W)-transformation (after bias correction)				
	binary	$\delta = 0.005$	$\delta = 0.001$	$\delta = 0.0005$	linear
$ln\,\mathcal{U}_{it}$	0.003	0.003	0.002	0.002	0.003
	(1.04)	(1)	(0.81)	(0.62)	(1)
$ln\,\mathcal{V}_{it}$	0.028***	0.028***	0.031***	0.032***	0.027***
	(4.23)	(4.27)	(4.74)	(4.87)	(4.16)
$ln\,A_{it}$	0.04	0.039	0.025	0.002	0.033
	(0.8)	(0.78)	(0.5)	(0.04)	(0.65)
λ	0.42***	0.428***	0.582***	0.701***	0.513***
	(14.22)	(15.55)	(17.01)	(17.29)	(15.16)
γ	0.583***	0.584***	0.57***	0.565***	0.582***
	(29.73)	(29.77)	(28.7)	(28.51)	(29.57)
δ	-0.202***	-0.215***	-0.272***	-0.301***	-0.241***
	(-5.34)	(-6.02)	(-6.09)	(-5.67)	(-5.55)
σ^2	0.004***	0.004***	0.004***	0.004***	0.004***
	(29.19)	(28.88)	(28.96)	(29.05)	(28.93)
log-like	2323.637	2339.839	2354.896	2357.996	2336.687
BIC	-2.953	-2.974	-2.993	-2.997	-2.970
obs.	1566	1566	1566	1566	1566

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010a). λ is the spatial autoregressive coefficient, γ captures the time dynamic effect and δ captures the combined spatial-time effect. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 1.10 Estimates of matching functions including the agglomeration index using the spatial dynamic panel model (2000-2009)

perform a robustness check to analyze whether cross-border commuting influences our results.[16] Therefore, we perform the same (monthly) regressions as before by leaving out those local employment offices that are located at the border.[17] After that reduction, our sample consists of 137 local employment offices.

For the sake of brevity, we only present the main features of the results.[18] The elasticities of matches on the unemployment stock in case of the basic and static model are smaller in comparison to the results using the full sample. The reduced sample covers a smaller labor market which implies in general less employment opportunities. Additionally, the decrease in the matching elasticity on unemployment might be an indication for substantial commuting of Germans to neighboring countries as residents

[16]We thank two anonymous referees to remind us of this issue.

[17]Note that we only exclude those local employment offices that are close to an actual border in contrast to those that border on the North Sea or Baltic Sea.

[18]The full results can be obtained from the author upon request.

that live farther away from the border are less likely to search in neighboring countries for a job. Regarding the matching elasticities on the vacancy stock and the spatial coefficients, there is no clear direction of change induced by the exclusion of border employment offices.

1.7 Conclusion

In this paper, we estimate German matching functions taking into account spatial dependencies. We show that German matching data exhibit significant spatial autocorrelation. To avoid biased and inefficient estimates, we apply a spatial econometric modeling to the matching function. Our panel data set covers monthly information for 176 local employment offices in Germany for the period from 2000 to 2009. In order to capture the dynamics on labor markets, we use not only a static modeling but also a dynamic model specification. For the estimation, we follow the methodology proposed in Lee and Yu (2010c) as well as in Lee and Yu (2010a) for the static and the dynamic model, respectively. We also introduce an agglomeration index as additional explanatory variable to control for the size of the labor market in neighboring regions. However, it is only significant in some of the specifications. In order to control for the impact of cross-border commuting, we perform a robustness exercise by excluding the border employment offices. To incorporate the spatial information into the model, we construct different spatial weights matrices. As the amount of commuting reflects interregional relations on labor markets, we exploit commuting data for the construction of different spatial weights matrices. Our results suggest that neglecting spatial dependencies yields upward-biased matching elasticities. Furthermore, they show that the dynamic model captures the structure in the data in a more appropriate way.

Regarding policy implications, our results suggest significant spatial spillovers. Accordingly, regional policy activities have wide consequences. On the one hand, a local unemployment shock is not limited to one region but has also effects on neighboring regions. But, on the other hand, regional activities aiming at a reduction of unemployment also have an impact on neighboring regions. Since we use numbers of commuters to measure the spatial impact in our model, neighboring regions are not limited to those that are neighbors in the literal sense. Hence, the presence and the range of spatial spillovers has to be taken into account when regional policy measures are designed.

Chapter 2

Explaining regional unemployment differences in Germany: a spatial panel data analysis

This chapter is based on Lottmann (2012a).

2.1 Introduction

The unemployment rate is a widely used and often discussed indicator for the economic well-being of a country. However, the discussion is mostly concentrated on national unemployment rates which give no information about the regional structure of unemployment. Though, data on regional unemployment rates show substantial differences between regions. According to Taylor and Bradley (1997), regional differences within a country are stronger than differences between countries. Due to the specific history of the country, regional differences are of particular interest in Germany. Until 1990, Germany was divided into two separate countries with different economic systems. The division of Germany caused structural differences resulting in adjustment processes which have not been fully completed until today. However, disparities are not only marked by structural differences between East and West Germany but there is considerable dispersion in regional unemployment across East and West German regions (Niebuhr et al. (2012)).

This paper analyzes determinants for regional differences in German unemployment rates using spatial econometric methods. We identify the driving factors in the whole of Germany as well as in East and West Germany separately. Twenty years after German reunification, this study is, to our best knowledge, the first contribution investigating regional unemployment in Germany.

A specific feature of regional labor markets is their correlation over space. The presence of spatial (auto-)correlation implies that the level of regional unemployment in one particular region is correlated with that of neighboring regions. On the one hand, firms do not restrict their recruiting activities to their resident location and, on the other hand, job searchers might accept a job in a different area. Patacchini and Zenou (2007) propose a model for this process to show that regional unemployment in some region depends positively on unemployment in neighboring regions in the previous period. The spatial econometric literature shows that ignoring spatial effects yields biased and inefficient estimates (see Anselin and Bera (1998) among others). Therefore, we apply a spatial econometric model to avoid these shortcomings.

To model regional unemployment, we take into account 24 possible explanatory variables containing equilibrium and disequilibrium factors and derive our set of regressors by a model selection procedure. We have panel data on 412 German districts for the period from 1999 until 2007. As labor market data exhibit not only spatial but also temporal dynamics, we utilize both a static and a dynamic modeling approach while most contributions in the literature consider only static model specifications.

Regional unemployment differentials have been subject of intensive research in the literature. From a methodological point of view, the empirical literature can be divided into two strands of literature. Firstly, models for regional unemployment are estimated using (non-spatial) panel data techniques. Examples are Partridge and Rickman (1997), who use data on state unemployment for the United States, as well as Taylor and Bradley (1997), who provide a comparative study for regional unemployment disparities in Germany, Italy and the United Kingdom. Their data for Germany covers only the Western part for the period from 1984 until 1994 and the data is on the level of German *Länder* which correspond to the NUTS I level.[1] Secondly, contributions apply spatial econometric models in a cross-sectional setting. The first contribution in this direction is by Molho (1995) in which he provides evidence for the presence of significant spillovers in the adjustment to local shocks using data on 280 local labor market areas in Great Britain. Further examples for this strand of literature are Aragon et al. (2003), who analyze district-level data for the Midi-Pyrénés region of France and Cracolici et al. (2007), who explore the geographical distribution of unemployment in Italy. Finally, Elhorst (2003) provides a survey on theoretical models and explanatory variables for regional unemployment differences.

We contribute to the existing literature by the following two aspects: Firstly, we apply both a static and a dynamic spatial panel model. Furthermore, we exploit the panel dimension of the data and, in addition to that, we account for both spatial and temporal dependence in the data. Our results show that the spatial dynamic panel model fits our data best. Secondly, we provide evidence that regional unemployment in Germany is of disequilibrium nature which provides a justification for political interventions on regional labor markets.

The structure of this paper is as follows: The second section briefly reviews theoretical explanations for regional unemployment differentials while the third presents the data set and explains how the spatial weights matrix is defined. The econometric model

[1]NUTS (French abbreviation) stands for "Nomenclature of Territorial Units for Statistics", and it is a hierarchical classification of regional units for statistical purposes

is introduced in the fourth section which covers model selection, specification testing and spatial econometric modeling. The fifth section is dedicated to the estimation results for the whole of Germany as well as for East and West Germany individually. Finally, the last section concludes.

2.2 Theoretical explanations for regional unemployment differentials

Classical economic theory suggests that differences in regional unemployment should not occur because unemployed living in a region with high unemployment are expected to move to an area with lower unemployment. A similar reasoning holds for firms which are assumed to move from low-unemployment to high-unemployment regions because they can benefit from a larger pool of workers. However, regional unemployment data shows substantial differences.

2.2.1 Why do regional unemployment rates differ?

The literature provides different explanations for the existence of regional unemployment differentials which can be summarized into two different views. The equilibrium view assumes the existence of a stable equilibrium in which regions have different unemployment rates. According to Molho (1995, p. 642), this equilibrium is characterized by "uniform utility across areas for (each) homogeneous labor group". In this setting, there is no incentive for further migration. Hence, households (and firms) need to be compensated for high (low) unemployment by other positive factors, so-called amenities. Such amenities are, for example, reasonable housing prices or higher quality of life. Therefore, the equilibrium rate of unemployment in region i is a function of the amenity endowment in this region (Marston (1985)). The equilibrium view has received theoretical and empirical support from Marston (1985) (among others) drawing on ideas from Hall (1970).

Contrary to the equilibrium view, the disequilibrium view assumes that regional unemployment will equalize in the long run. However, the adjustment process might be slow. The speed of adjustment depends on different factors that are connected to both labor supply and demand. Such factors are, for example, the age structure and educational attainment of the population. Young people are more likely to migrate as they have lower opportunity costs and are less risk averse (Aragon et al. (2003)). People holding a degree of higher education are also more likely to move because the labor market for high-skilled workers is larger and these individuals are expected to be better informed (Aragon et al. (2003)). The structure of the labor force also influences the relocation behavior of firms. Moreover, population density also affects the adjustment process to the long-run equilibrium. Unemployment is expected to be lower in urban areas because the matching process between unemployed and vacant jobs is more efficient. Furthermore, the migration behavior of people is clearly influenced by migration costs. For example, housing prices and the structure of the housing market influence how easy it is for a household to change its location.

These explanations for regional unemployment differences give rise to different conclusions for policy makers. According to Marston (1985, p. 58), government efforts to reduce regional unemployment differentials are "useless" since they cannot reduce unemployment anywhere in the long run when the level of regional unemployment can be considered as equilibrium state. By contrast, the disequilibrium view delivers an "implicit justification for programs that target government funds to depressed areas" (Marston (1985), p. 58). In light of these different consequences for policy, it is important to assess whether regional unemployment can be considered as equilibrium phenomenon or not.

However, both explanatory approaches for regional unemployment are not necessarily mutually exclusive. Marston (1985, p. 59) states that "it may be that an equilibrium relationship exists, but that equilibrating forces are so weak that individual areas spend a long period of time away from their equilibrium". For the German case, there are arguments for both theoretical approaches to explain the regional labor market situation. On the one hand, about twenty years after German reunification, the economic catching-up process of East Germany is not yet complete. But, on the other hand, regional unemployment rates are not expected to equalize in the long run because of structural differences between regions. Structural differences exist not only between but also within East and West Germany and other parts.

Partridge and Rickman (1997) combine both approaches and extend the set of factors that might influence regional unemployment. In contrast to the equilibrium view, they do not assume that household utility in terms of income and amenities will equalize across areas in equilibrium. They add monetary and psychological costs of household relocation to the household utility function. These costs can be sufficiently high such that moving of households is limited. As regional unemployment in Germany has both equilibrium and disequilibrium aspects, we base our empirical analysis on Partridge and Rickman (1997).

2.2.2 Set of possible determining factors

Following Partridge and Rickman (1997), we assume that unemployment in region i in year t depends on disequilibrium variables and an equilibrium component which is a function of market equilibrium effects, demographic characteristics as well as producer and consumer amenities. For the choice of the actual variables in these categories we take into account the empirical regional unemployment literature. However, the set of our variables is limited by data availability.

Disequilibrium effects We use the employment growth rate which, according to the literature, has turned out to be an important determinant for regional unemployment. This is not surprising because the change in employment directly affects unemployment.[2] Another variable capturing disequilibrium effects are wages or unit labor costs. Unfortunately, this data is not available for our study period.

Market equilibrium effects To account for the sectoral structure of regions, we use employment shares of different sectors. According to Martin (1997, p. 244), industrial composition effects are a "primary reason" for labor demand and regional unemployment to differ across regions.

Demographic characteristics Demographic characteristics influence both labor demand and labor supply by affecting the number of new hires, quits and workers leaving the labor force (Partridge and Rickman (1997)). We use the share of young and older workers to account for the age structure of the labor force. In contrast to studies on other countries, as for example the United States, German labor market data do not contain any information on ethnicity in general. However, we have data on the share of foreigners in the labor force. Another important demographic variable is labor force participation, especially female labor force participation. Due to different social roles of women in both German countries before 1990, labor force participation of women differs substantially between East and West Germany. Unfortunately, the data on female labor force participation is only available on the level of *Regierungsbezirke* which partly correspond to the NUTS II regions of Germany, i.e. this variable exhibits less regional variation than the others. To include information on human capital, we utilize data providing three levels of educational attainment which are a university degree, a vocational qualification and no professional qualification at all. Furthermore, we use the balance of incoming and outgoing commuters of district i to control for a region's linkages with other regions. A positive commuting balance in region i indicates that labor supply in region i increases by incoming commuters. Moreover, a positive commuting balance gives an indication for labor demand exceeding labor supply in region i.

Amenities On the one hand, the impact of amenities is captured by population density. It is a proxy for consumer and producer amenities because urban areas provide more amenities than rural areas. Jobless individuals have more employment opportunities and the matching process is expected to be more efficient in urban areas. However, urban areas are also associated with pollution and congestion. On the other hand, we consider three amenity variables which, to our best knowledge, have not been considered in the regional unemployment literature so far. First, we use the public debt ratio of a district because high public debts in relation to gross domestic product (GDP) are an indication for a deficient ability of a region to finance public goods and subsidies. Additionally, strongly indebted communities are not attractive for firms to create new

[2] It would be interesting to analyze the impact of (temporally) lagged values of employment growth on regional unemployment. However, to our best knowledge, employment data on periods prior to 1999 is not available on the level of districts.

businesses. Second, we utilize data on the number of business registrations. This variable is a proxy for producer amenities. A higher number of new businesses will result in a higher demand for labor. Third, we use the number of overnight stays to capture a region's attractiveness to tourists. Additionally, a high number of overnight stays may be related to high business activities.[3] Note that our amenity variables differ across regions and years. Hence, they are not differenced out in a fixed effects specification.

2.3 Data and spatial weights matrix

2.3.1 Regional unemployment and its determining factors

The data on regional unemployment rates used in this analysis are provided online by the Federal Employment Office (*Bundesagentur für Arbeit*). As it is official data, the underlying definition of unemployment corresponds to regulations in German Social Security Code (*Sozialgesetzbuch*). Moreover, we utilize a huge regional data set of possible explanatory variables. All these variables are taken from the regional database of the Federal Statistical Office of Germany (*Statistisches Bundesamt*). Since there were some values missing in this database, we requested them directly from the corresponding regional statistical institutions. A detailed description of the data and sources can be found in Table B.3 in Appendix B. Our data set covers the period from 1999 until 2007.[4] The end of our sample period is determined by a change in the sectoral classification in 2008, i.e. data on employment in different industries is not comparable before and after this change of classification. The data is available for all 412 German districts (*Landkreise* and *kreisfreie Städte*) which correspond to German NUTS III regions.[5] During our sample period, there are two reforms of district allocation. We allocate the data for the whole period in such a way that it corresponds to the situation after these reforms. Details on the district reforms can be found in Appendix B.

To visualize regional differences in unemployment rates of German districts, Figure 2.1 presents a map of Germany which is colored according to the extent of regional unemployment in 2009.[6] Additionally, Table 2.3 shows summary statistics of regional unemployment rates over time. Based on these exploratory tools, we can summarize

[3] In contrast to other studies (as Cracolici et al. (2007) or Molho (1995)), we do not consider housing prices in our analysis because the majority of Germans lives in rented apartments. In 2006, 58% of the German population lived in rented appartments (see Timm (2008)). Until now, there exists no comprehensive data base for rental prices in German districts.

[4] In 2005, a labor market reform ("*Hartz* reform") became effective which changed the definition of unemployment. Therefore, the number of unemployed increased by definition in this year.

[5] Baddeley et al. (1998, p. 204) state that NUTS III regions "most closely approximate meaningful labor markets". However, Eckey et al. (2007) explain that travel-to-work areas are the relevant regional level for analyses of regional production and unemployment.

[6] The map of Germany shows that some of the NUTS III regions lie within others, i.e. these districts have only one physical neighbor.

the following major facts: First, there is substantial variation in regional unemployment rates in Germany. In 2004, the district with lowest unemployment exhibited a rate of 4.4 % (Eichstätt district) while the highest regional unemployment rate amounted to 31.4 % (Uecker-Randow district). Second, the German labor market is characterized by strong differences between East and West Germany which can be considered as consequences of the former German division. Regional unemployment rates are higher in East Germany. However, in a ranking of German districts with respect to unemployment, not all East German districts are placed behind the districts of West Germany. Third, besides the East-West differences, there is a slight North-South divide in regional labor market performance. These findings are in line with those of Niebuhr et al. (2012) who state that disparities in regional labor market performance are substantial and have partly been increasing over recent years.

To test for stationarity of the data, we apply panel unit root tests. The results of the Im et al. (2003) (IPS) test and the Fisher-type (ADF) test, which was proposed in Maddala and Wu (1999) and in Choi (2001), clearly reject the hypothesis of a unit root in regional unemployment rates at all reasonable significance levels. In addition to that, we apply the IPS test and the Fisher-type (ADF) test to our set of explanatory variables and also find that all explanatory variables are stationary. However, Baltagi et al. (2007a) show that there can be considerable size distortions in panel unit root tests when the true model exhibits spatial error correlation. Hence, these test results can only serve as a slight indication regarding stationarity of the data.

2.3.2 Spatial autocorrelation on German labor markets

An important component of spatial econometric models is the spatial weights matrix. It is a nonstochastic matrix that specifies exogenously the spatial relations between observations. Hence, the spatial weights matrix determines the neighborhood of each district. Accordingly, the term 'neighboring' always refers to the neighborhood set defined by the corresponding spatial weights matrix. We use both a binary spatial weights matrix with entries zero and one and matrices with general weights.

The simplest version of a spatial weights matrix is the binary contiguity matrix. When two districts share a common border, the corresponding entry in the spatial weights matrix is one and zero otherwise. The elements on the main diagonal are zero by definition. This matrix induces a simple spatial structure which might not reflect actual spatial linkages in an appropriate way. Therefore, we construct spatial weights matrices with general weights. On the one hand, we utilize data on geographic distances between districts and, on the other hand, we use a combination of geographic distance and size, as proposed in Molho (1995), to define spatial weights.

Geographic distance has frictional effects on labor market activity. Workers prefer to find a job in their closer environment because commuting and moving entail monetary and psychological costs. Therefore, we use great circle distances between centroids of districts to define the entries of the spatial weights matrix. Summary statistics of the geographic distances are provided in Table 2.1.

Min	1st Qu.	Median	Mean	3rd Qu.	Max	Std. dev.
1.18	191.7	298	310.6	417.1	845.6	155.52

Table 2.1 Summary statistics of geographic distances (in kilometers) between centroids of German districts

The weights of the distance-based matrix are defined by

$$w_{ij} = \begin{cases} exp(-\tau d_{ij}) & \text{for } i \neq j \\ 0 & \text{for } i = j, \end{cases} \tag{2.1}$$

where τ is a distance decay parameter and d_{ij} is the geographic distance between districts i and j. The resulting spatial weights matrix crucially depends on the choice of τ. To determine the distance decay parameter, we use a grid search with different values for τ and decide according to the Bayesian and Akaike's information criterion which parameter value is most suitable for our data. Niebuhr (2003) also uses this distance decay function to define the weights for her analysis of regional unemployment in Europe.

However, the distance decay function neglects the labor market size of districts. Spatial dependence differs when the extent of employment opportunities differs although distances between districts are the same. We expect that the spatial impact of a district with high employment on a low-employment district is stronger than vice versa. Therefore, we utilize the weighting scheme proposed by Molho (1995) which combines size with the distance decay effect. According to Molho (1995), the spatial weights are defined by

$$w_{ij} = \begin{cases} \frac{E_j\, exp(-\eta d_{ij})}{\sum_{k \neq i} E_k\, exp(-\eta d_{ik})} & for\ i \neq j \\ 0 & for\ i = j, \end{cases} \tag{2.2}$$

where E denotes the employment level and η is the distance decay parameter. As Molho (1995) points out, this weighting scheme implies that the spillover effect of the labor market situation in region j on the setting in region i increases with size of region j (measured in terms of employment) and decreases with the distance between both districts. Again, the impact of distance on the strength of the spatial relation crucially depends on the distance decay parameter η. We perform a grid search for η and decide on the appropriate value for our model according to information criteria.

Labor market activity and hence labor market data is expected to be correlated over space. To justify this aspect, we perform the Moran I test for spatial autocorrelation using regional unemployment rates. As this test is not specified for a particular spatial process, we can apply it directly to our data. The null hypothesis of this test is the absence of spatial autocorrelation while the alternative is not exactly specified. The test

	Moran I	Z	p-value
1999	0.874	26.48	0
2000	0.875	29.02	0
2001	0.890	29.51	0
2002	0.882	29.25	0
2003	0.863	28.61	0
2004	0.846	28.05	0
2005	0.799	26.5	0
2006	0.810	26.86	0
2007	0.793	26.29	0

Notes: Z denotes the standard deviate of the Moran I statistic, i.e. $Z = \frac{I-E[I]}{sd(I)}$. The null hypothesis is the absence of spatial autocorrelation whereas the alternative is positive spatial autocorrelation. The Moran I values are computed assuming normality.

Table 2.2 Results of the Moran I test for spatial autocorrelation (1999-2007)

statistic can be expressed by (Moran (1950))

$$I = \frac{\sum_i^n \sum_j^n w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_{i=1}^n (u_i - \bar{u})^2} \tag{2.3}$$

where u_i and u_j are the regional levels of unemployment in district i and j. $\bar{u}$ is defined by $\bar{u} = \frac{1}{n}\sum_{i=1}^n u_i$ and w_{ij} is the element of the spatial weights matrix indicating the spatial impact of region j on region i. For the computation of the Moran I statistic we use the binary contiguity matrix.[7]

As the Moran I statistic is designed to detect spatial autocorrelation in cross-sectional data, we compute it for every year of our sample separately. The results of the Moran I test are presented in Table 2.2. They show that regional unemployment rates are positively spatially autocorrelated during the period from 1999 until 2007. Furthermore, they show a decreasing trend in the values of the Moran I statistic, i.e. the extent of spatial autocorrelation in regional unemployment rates decreases during 1999 and 2007.

2.4 Econometric Model

In order to control for spatial autocorrelation in the data, we specify a spatial econometric model for our analysis of regional unemployment. We apply a panel data model which allows to account for unobserved individual heterogeneity in the data. We obtain our model in two steps: Firstly, we use a model selection procedure to decide which

[7]We also tried the other spatial weights matrix to compute the Moran I statistic and got qualitatively the same results.

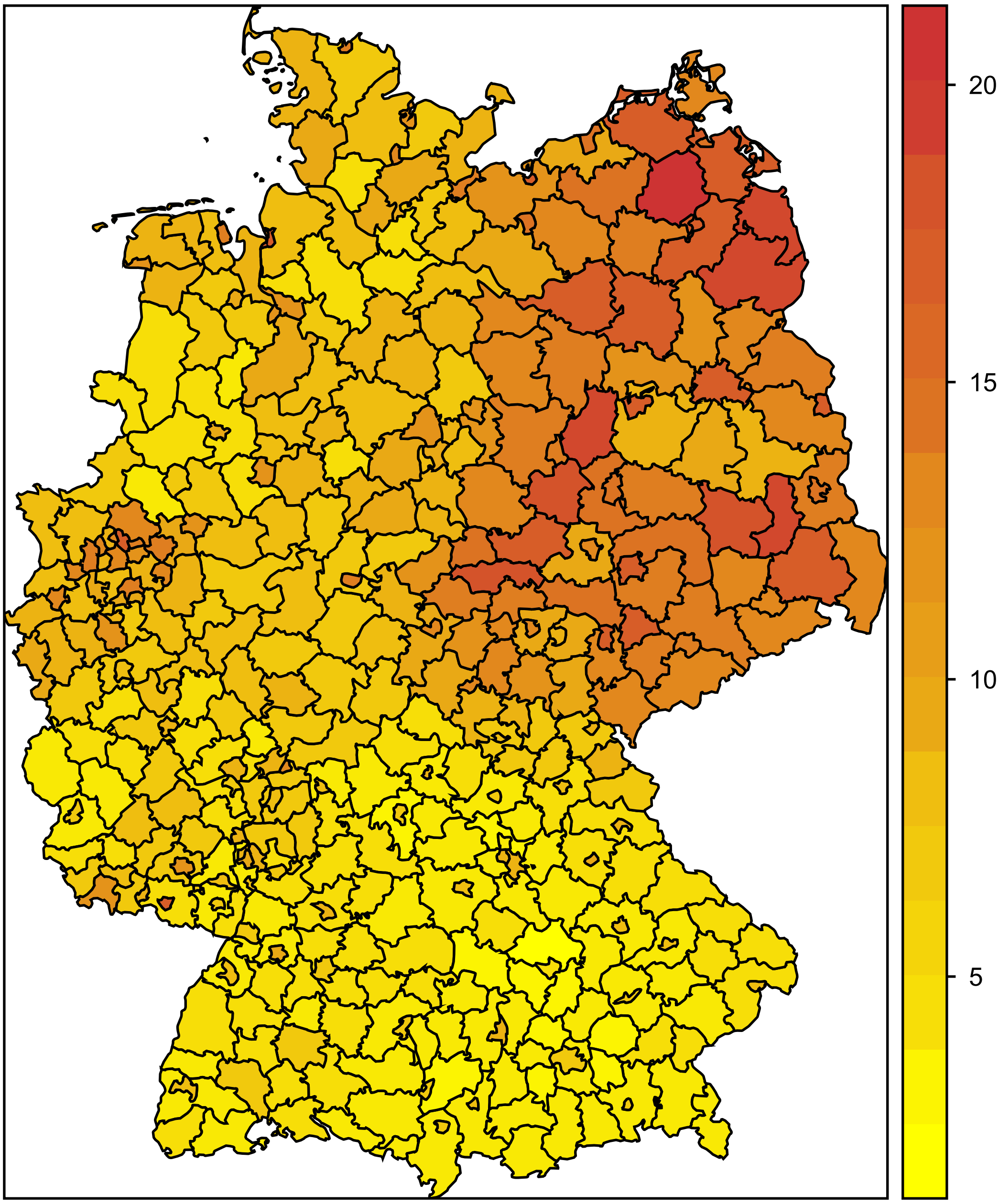


Figure 2.1 Regional unemployment in Germany in 2009

	Min	1st Qu.	Median	Mean	3rd Qu.	Max	Std.dev.	national
1999	4	7.8	10	11.41	14.3	24.8	4.815	11.7
2000	3	6.7	8.8	10.46	13.3	25.6	5.158	10.7
2001	3	6.3	8.4	10.19	12.7	26.7	5.356	10.3
2002	3.9	6.9	9	10.69	12.9	27.6	5.279	10.8
2003	4.6	7.7	9.8	11.57	13.9	29.7	5.424	11.6
2004	4.4	7.7	9.8	11.66	14	31.4	5.467	11.7
2005	4.7	8.7	11.4	12.84	16.1	29.7	5.323	13
2006	3.7	7.7	10.5	11.81	15	27.6	5.084	12
2007	2.4	6.1	8.5	9.868	12.6	24.2	4.733	10.1
2008	1.9	4.8	7.2	8.435	11	21.5	4.306	8.7
2009	2.5	5.7	7.9	8.843	11.4	20.1	3.908	9.1

Table 2.3 Summary statistics of regional unemployment rates (1999-2009)

variables from our set of possible explanatory variables actually have a significant impact on regional unemployment. Secondly, we use the specification test by Debarsy and Ertur (2010) to assess which spatial process captures the spatial dynamics in our data in the best way.

2.4.1 Model selection

Our model selection procedure is based on the standard two-way fixed effects panel model (Baltagi (2008)), i.e.

$$u_{it} = \sum_{k=1}^{K} \beta_k x_{kit} + \mu_i + \alpha_t + \epsilon_{it}; \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{2.4}$$

where u_{it} is the regional unemployment rate, β_k are unknown parameters and x_{kit} are the values of K explanatory variables. μ_i denotes district-specific effects and α_t represent time effects. We assume the district-specific effects to be fixed as our data set contains information on all German districts. The time effects capture national factors as, for example, business cycle effects that affect all regions in the same way. ϵ_{it} are the disturbances for which it is assumed that $\epsilon_{it} \sim (0, \sigma_\epsilon^2) i.i.d.$ The indices of the variables denote district i and year t.

Model (2.4) controls neither for spatial autocorrelation nor for temporal dynamics in the data. Therefore, we refer to this model as basic model. If spatial dependence in the data is ignored, standard OLS regression will provide biased parameter estimates in case of spatial lag dependence and in case of spatially lagged exogenous variables. However, OLS estimation produces unbiased but inefficient estimates for the spatial error model. Neglecting a spatial lag term is similar to an omitted variable bias (Franzese and Hays (2007)). As the spatial lag term is correlated with the error term, OLS estimation of the associated coefficient will be inconsistent (Franzese and Hays (2007), Anselin and Bera (1998)).

In order to choose the relavant variables, we divide our set of explanatory variables into three groups according to theoretical importance. Then, we regress regional unemployment rates on different combinations of variables where the variables with the strongest theoretical support are always contained. To keep compuational effort manageable, we base these regressions on the basic model (equation (2.4)), although OLS estimation produces biased and/or inefficient results for spatially autocorrelated data. Finally, we compute Akaike's (AIC) and the Bayesian information criterion (BIC) to assess the goodness-of-fit of the regressions.

Table 2.4 provides an overview of the division of explanatory variables into these groups. The first group of variables contains variables which are essential for our model. We include into this group the employment share in manufacturing and in the construction industry ($\%IND$ and $\%CON$), the age-related demographic variables ($YOUNG$ and OLD) as well as one of the human capital variables ($H0$). Additionally, we include employment growth (EG) in this group to account for disequilibrium effects.[8] The second group contains variables that are expected to be important for the explanation of regional unemployment rates. We assign to this group our amenity variables ($DENS$, $DEBTR$, $STAY$ and REG). Furthermore, we consider the employment shares of agriculture ($\%AGR$), electricity, gas and water supply ($\%ENERW$), financial business ($\%FIN$), transport, storage and communication ($\%TRANS$), real estate ($\%REAL$) and public administration ($\%PUB$) in this group. Additionally, this group contains female labor force participation (FP) as well as the remaining educational variables ($H1$ and $H2$). The last group consists of variables that are expected to have a weaker influence on regional unemployment. These variables are the share of foreign employed individuals ($FOREIGN$) and the employment shares of mining and quarrying ($\%MINE$), wholesale and retail trade ($\%TRADE$), hotels and restaurants ($\%HOT$) as well as education, health and social work ($\%EDUHEALTH$).

Our model selection procedure selects a model containing thirteen variables. The summary statistics of these variables are in Table B.1 in Appendix B. To check for possible multicollinearity in our model, we analyze both the correlation matrix of the regressors and variance inflation factors where both give no indication for multicollinearity. Hence, our final best model is

$$\begin{aligned}
u_{it} = {} & \beta_1 EG_{it} + \beta_2 \%IND_{it} + \beta_3 \%ENERW_{it} + \beta_4 \%CON_{it} + \beta_5 \%HOT_{it} \\
& + \beta_6 \%FIN_{it} + \beta_7 \%PUB_{it} + \beta_8 YOUNG_{it} + \beta_9 OLD_{it} + \beta_{10} H0_{it} \\
& + \beta_{10} H1_{it} + \beta_{12} REG_{it} + \beta_{13} DEBTR_{it} + \mu_i + \alpha_t + \epsilon_{it}; \\
& i = 1, \ldots, n; \quad t = 1, \ldots, T, \qquad (2.5)
\end{aligned}$$

[8]Note that we have not assigned female labor force participation to this group as its regional variation is small because of limited data availability.

where the variables are defined as before. The time effects (α_t) are strongly correlated with the national unemployment rate (correlation: 0.95).

Our final model contains all variables of group one. The model selection procedure selects the share of employed individuals holding a vocational qualification as additional demographic variable. Hence, we account for two of three educational variables. Only the public debt ratio and the number of business registrations of our amenity variables are contained in our model. Thus, our model selection results reveal a first slight indication that regional unemployment is a disequilibrium phenomenon. Furthermore, the age-related demographic variables and the educational variables are contained in our final model. Regarding the market equilibrium effects, employment shares in electricity, gas and water supply, hotels and restaurants, financial business and public administration are selected into our model in addition to the sectoral variables of group one. The significance of the employment share in hotels and restaurants can be explained by the fact that a significant part of the work in this industry is done by workers holding no specific training qualification for this field. Hence, it might be easier for jobless individuals to find a job in this field.

group 1	group 2	group 3
- employment growth (EG)	- female labor force participation (FP)	- share of foreign employed individuals ($FOREIGN$)
share of individuals working	share of employed individuals	share of individuals working
- in manufacturing ($\%IND$)	- with vocational training ($H1$)	- in mining and quarrying ($\%MINE$),
- in construction industry ($\%CON$)	- and with university degree ($H2$)	- in hotels and restaurants ($\%HOT$),
share of	- population density ($DENS$)	- in wholesale and retail trade ($\%TRADE$),
- young ($YOUNG$)	- public debt ratio ($DEBTR$)	- in education, health and social work ($\%EDUHEALTH$)
- old workers (OLD)	- business registrations (REG)	
- employed individuals without any vocational training ($H0$)	- number of overnight stays ($STAY$)	
	share of individuals working	
	- in agriculture, hunting and forestry ($\%AGR$),	
	- in electricity, gas and water supply ($\%ENERW$),	
	- in transport, storage and communication ($\%TRANS$),	
	- in financial business ($\%FIN$),	
	- in real estate, renting and business activities ($\%REAL$),	
	- in public administration and defence; compulsory social security ($\%PUB$)	

Table 2.4 Division of explanatory variables for model selection

2.4.2 Spatial econometric modeling

To capture the spatial dependence in the data, we specify a spatial panel model. The spatial econometric literature provides different models for data with spatial autocorrelation: the model with spatially lagged exogenous variables (SLX model), the spatial error model, the spatial lag model and combinations of them. The SLX model is, from a methodological perspective, the simplest model because the additional regressors are exogenous and the error term remains spherical. We estimated this model for our data. Although the results are slightly better (according to information criteria) than those of the basic model, we find significant spatial dependence left in the residuals based on the results of the test by Pesaran (2004). Moreover, the results of the SLX model are worse compared to those of our more complex spatial models.[9]

Testing for the spatial model specification

As the model with spatially lagged exogenous variables is not appropriate for our data, we need to specify one of the other spatial processes. Hence, we perform the specification test by Debarsy and Ertur (2010) to differentiate between the spatial models. To our best knowledge, the test by Debarsy and Ertur (2010) is the only specification test that allows to discriminate between the spatial lag model, the spatial error model and the model including both a spatial lag and spatially autocorrelated errors. Baltagi et al. (2003) extend the langrange multiplier (LM) test by Breusch and Pagan (1980) to the spatial error component model to test simultaneously for the existence of spatial error correlation as well as for random region effects. Additionally, they derive conditional tests for spatial error correlation and random region effects. Baltagi et al. (2007b) generalize the underlying model to a spatial panel model that controls for serial correlation over time for each spatial unit. We use this test to motivate our spatial dynamic model. Finally, Baltagi and Liu (2008) derive a test for autoregressive spatial lag dependence instead of spatial error terms.

The starting point of the test by Debarsy and Ertur (2010) is the spatial autoregressive model with spatially autocorrelated disturbances of order $(1,1)$ (SARAR (1,1) model), i.e.

$$U_t = \lambda W U_t + X_t \beta + \mu + V_t; \quad V_t = \rho W V_t + \Xi_t; \quad t = 1, \ldots, T, \tag{2.6}$$

where $U_t = (u_{1,t}, u_{2,t} \ldots, u_{n,t})'$ is a $(n \times 1)$ vector containing regional unemployment rates. X_t is the $(n \times k)$ matrix containing all explanatory variables from our selected model (equation (2.5)), β is the $(k \times 1)$ coefficient vector and $\mu = (\mu_1, \ldots, \mu_N)'$. W is the $(n \times n)$ spatial weights matrix.[10] $\Xi_t = (\xi_{1,t}, \ldots, \xi_{n,t})'$ is the $(n \times 1)$ vector of innovations

[9] The results can be obtained from the author upon request.

[10] Debarsy and Ertur (2010) specify the model in their original contribution using different spatial weights matrices for the spatial lag and spatial error part. But they note that the test also works when these are equal.

	H_0^a	H_0^b	H_0^c	H_0^d	H_0^e
LM	1353.8	1285.7	967.19	7.86	3771.1
p-value	0	0	0	0.0051	0

Table 2.5 Test results of the specification test by Debarsy and Ertur (2010) using the binary contiguity matrix

where $\xi_{i,t}$ are i.i.d. across i and t and $\xi_{i,t} \sim (0, \sigma_\xi^2)$. Finally, λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient.

Debarsy and Ertur (2010) consider five different hypotheses in their paper:

- $H_0^a : \rho = \lambda = 0$. This joint hypothesis tests whether there is spatial dependence in the data at all. If it cannot be rejected, there is no need for a spatial econometric model.
- $H_0^b : \lambda = 0$. Under the alternative, the specification is the spatial lag model. However, spatial errors may exist.
- $H_0^c : \rho = 0$. Under the alternative, the model contains spatially autocorrelated errors. However, a spatial lag term may exist.
- $H_0^d : \rho = 0$, with λ possibly different from 0. Under the alternative, the general specification (equation 2.6) has to be estimated.
- $H_0^e : \lambda = 0$, with ρ possibly different from 0. Under the alternative, the general specification (equation 2.6) has to be estimated.

The test statistics for the hypotheses H_0^a until H_0^e are in Appendix B. Table 2.5 shows the results of the Debarsy/Ertur (2010) test using the binary contiguity matrix.[11] According to the results, we can reject all five hypotheses even on the 1% significance level. Hence, the SARAR(1,1) model is the most appropriate model for our data.

Static model specification

In accordance with the results of the test by Debarsy and Ertur (2010), we include a spatial lag term and spatially autocorrelated errors in our model. Additionally, we incorporate time effects in our static spatial panel model in order to have a two-way specification as in our basic model. The static model specification is

$$\begin{aligned} U_t = {} & \lambda WU_t + \beta_1 EG_t + \beta_2 \%IND_t + \beta_3 \%ENERW_t + \beta_4 \%CON_t + \beta_5 \%HOT_t \\ & + \beta_6 \%FIN_t + \beta_7 \%PUB_t + \beta_8 YOUNG_t + \beta_9 OLD_t + \beta_{10} H0_t + \beta_{11} H1_t \\ & + \beta_{12} REG_t + \beta_{13} DEBTR_t + \mu + \alpha_t \mathbb{1}_n + V_t; \quad V_t = \rho WV_t + \Xi_t; \quad t = 1, \ldots, T, \end{aligned} \tag{2.7}$$

[11]We also performed this test using the other spatial weights matrices and obtained qualitatively the same results.

where the variables are defined as before. The elements of the $(n \times 1)$ disturbance vector $\Xi_t = (\xi_{1,t}, \ldots, \xi_{n,t})'$ are assumed to be i.i.d. across i and t with zero mean and constant variance σ^2_{ξ}. $\mathbb{1}_n$ denotes a $(n \times 1)$ vector of ones.

Lee and Yu (2010c) show that for the (static) model with fixed individual and time effects the direct quasi-maximum likelihood estimation method yields inconsistent estimates for the common parameters unless n is large. In addition to that, they show that even in the case when both n and T are large, the distribution of the estimates of common parameters is not properly centered.

Moreover, Lee and Yu (2010c) show that the use of the typical within transformation to eliminate fixed effects causes the errors in the within-transformed model to be linearly dependent. Therefore, they apply an orthogonal transformation to eliminate the individual effects which produces independent error terms. The standard within transformation uses the deviation from time mean operator, i.e. $J_T = I_T - \frac{1}{T}\mathbb{1}_T\mathbb{1}_T'$ where I_T is the identity matrix of dimension T. Lee and Yu (2010c) define the orthonormal eigenvector matrix of J_T, i.e. $[F_{T,T-1}, \frac{1}{\sqrt{T}}\mathbb{1}_T]$. $F_{T,T-1}$ is the $(T \times (T-1))$ submatrix corresponding to the eigenvalues of one. They suggest to transform the original data by $F_{T,T-1}$, i.e.

$$[Y^*_{n1}, \ldots, Y^*_{n,T-1}] = [Y_{n1}, \ldots, Y_{nT}]F_{T,T-1}. \tag{2.8}$$

Note that the dimension of the transformed model is $n(T-1)$. To remove the time effects from the model, they propose a similar transformation which is based on the orthogonal transformation using $J_n = I_n = \frac{1}{n}\mathbb{1}_n\mathbb{1}_n'$. Correspondingly, the model has dimension $(n-1)(T-1)$ after both transformations. Lee and Yu apply this transformation approach in various contributions (Lee and Yu (2010b), Lee and Yu (2010c), Lee and Yu (2010a)). We apply it to both our static and our dynamic model. Finally, the transformed model can be estimated by quasi-maximum likelihood.[12]

Dynamic model specification

Labor market data is not only correlated over space but also over time. To motivate the dynamic approach, we use the test by Baltagi et al. (2007b) because it allows for serial correlation in the error terms (in addition to spatial autocorrelation). Details on hypotheses and test statistics are Appendix B. The test results clearly show the following three aspects of our data. Firstly, there is serial dependence in our data. Hence, a dynamic model specification is reasonable in our context. Secondly, the test results give an indication for the presence of spatially autocorrelated errors. This is in line with the results of the Moran I test that also show significant spatial autocorrelation in regional unemployment rates. Thirdly, the test results support our assumption of a fixed effects model because we cannot reject the hypothesis that the standard deviation of the fixed effects is equal to zero.

[12] For more details on the estimation methodology, see Lee and Yu (2010c).

The literature on spatial dynamic panel models provides various model specifications. Elhorst (2012) provides a survey of the literature on specification and estimation of spatial dynamic panel data models. For our analysis of regional unemployment, we include a spatial lag term, a temporally lagged term as well as a combined spatially and temporally lagged term in our dynamic model. The resulting model can be described by

$$\begin{aligned} U_t = \lambda WU_t + \gamma U_{t-1} + \delta WU_{t-1} + \beta_1 EG_t + \beta_2 \%IND_t + \beta_3 \%ENERW_t + \beta_4 \%CON_t \\ + \beta_5 \%HOT_t + \beta_6 \%FIN_t + \beta_7 \%PUB_t + \beta_8 YOUNG_t + \beta_9 OLD_t \\ + \beta_{10} H0_t + \beta_{11} H1_t + \beta_{12} REG_t + \beta_{13} DEBTR_t + \mu + \alpha_t \mathbb{1}_n + \Xi_t; \quad t = 1, \ldots, T, \end{aligned} \tag{2.9}$$

where γ captures the pure time-dynamic effects and δ captures the combined spatial-temporal effect. The assumptions about the error term Ξ_t are as before.

Yu et al. (2008) propose a bias corrected quasi-maximum likelihood estimator for the spatial dynamic panel data model including a spatial lag, a temporal lag and a combined spatial-temporal term. However, they only allow for individual-specific fixed effects but not for fixed time effects. Lee and Yu (2010b) provide an estimator for the same model but extended to include time period fixed effects. Lee and Yu (2010b) show that direct quasi-maximum likelihood estimation of all parameters in the model with time effects yields an additional bias of order $O(n^{-1})$. They apply their transformation approach and show that it can avoid the additional bias with the same asymptotic efficiency as the direct quasi-maximum likelihood estimates when n is not relatively smaller than T. Furthermore, Lee and Yu (2010b) show that the direct estimates have a degenerate limit distribution while the transformed estimates are properly centered and asymptotically normal. Therefore, we apply the estimation methodology of Lee and Yu (2010b) to our dynamic model.

2.5 Estimation results

Firstly, we estimate the basic model, i.e. the model without any terms controlling for spatial or temporal dependence. The basic model is specified according to a two-way fixed effects panel data model and it is estimated using the standard within-estimator (see Baltagi (2008)). Secondly, we estimate the static spatial panel specification and, thirdly, the spatial dynamic model, both using the binary contiguity matrix, the distance decay matrix as well as the Molho (1995) weights matrix. Hence, we perform seven regressions for the whole of Germany. The regression results for the basic and the static model are in Table 2.6 and the results for the dynamic model are in Table 2.7. In addition to that, we perform the same regressions for the Eastern and Western part of Germany individually. Elhorst (2012) discusses stationarity issues and proposes stationarity conditions for spatial dynamic panel data models. These conditions as well as the conditions stated in Lee and Yu (2010a) are satisfied in the regression results for the whole of Germany. However, the regression results for East and West Germany using the distance decay matrix do not meet the stationarity conditions. Therefore, we only present the results using the other spatial weights matrices for the separate analyses.

2.5.1 Results for the whole of Germany

Economic interpretation As expected, regional unemployment rates are influenced negatively by employment growth. Furthermore, the shares of employed individuals working in manufacturing and in the construction industry also have a negative impact on regional unemployment. Hence, districts that are specialized in these industries exhibit lower unemployment than districts with a different sectoral structure. Our estimation results reveal no indication for a discrimination of older workers as the associated coefficient is also negative. Though, this coefficient should not be overinterpreted because it can simply be related to effects of demographic change, i.e. an aging labor force. By contrast, the impact of younger employees on regional unemployment is positive. But this does not impliy necessarily youth unemployment because the majority of people aged between 15 and 25 is still in the educational system. The share of employed individuals without any professional qualification influences regional unemployment positively which is in line with expectation from theory. Interestingly, this also holds for the share of employed individuals with vocational training.

Our model contains only a few of the amenity variables. Additionally, the signs of the amenity variables are against expectation from theory. According to the equilibrium view, consumers are expected to stay in regions with high unemployment when this region offers a great extent of amenities. Hence, high unemployment should be related negatively to public debt because heavily indebted districts are not able to finance public goods to improve life quality. If high public debts result from high investments in the past, consumers expect less expenditures in the future. However, our results show a significant positive coefficient for the public debt ratio. A similar reasoning holds for producer amenities. Firms are expected to move to districts with high unemployment, i.e. the level of producer amenities should be higher when regional unemployment is lower. But the coefficient of business registrations is positive in our empirical results. Even if the public debt ratio is interpreted as a proxy for producer amenities, its coefficient does not have the desired sign. Thus, our results reveal no indication for regional unemployment to be of equilibrium nature in Germany. Nonetheless, some of the market equilibrium variables, i.e. employment shares, are significant in our model.

Spatial econometric interpretation Ignoring spatial dependence in the data, results in biased and inefficient estimates. The estimated coefficients of the basic model are mostly upward-biased in absolute value in comparison with the results of the static model. In an earlier contribution (Lottmann (2012b)) we get a similar result for the estimation of matching functions. The existence of this bias is theoretically shown in Franzese and Hays (2007). In addition to that, the information criteria show that the spatial models are more appropriate for our data than the basic one. Hence, a spatial model is needed for the analysis of regional unemployment.

The dynamic model fits our data better than the static model according to information criteria. Thus, in order to model regional unemployment, a dynamic modeling approach needs to be applied. To our best knowledge, most of the contributions to the regional unemployment literature apply only a static model. However, most of the explanatory variables are not significant in the dynamic model. Hence, the temporal lag is able to explain a lot of the variability in regional unemployment rates. Only employment growth, the employment shares of manufacturing, construction industry and electricity, gas and water supply as well as the age-related demographic variables have a significant impact on regional unemployment.

The spatial autoregressive (λ) and the spatial autocorrelation coefficient (ρ) measuring the spatial influence in our static spatial panel model are both significant while the influence of both coefficients is positive in most cases. Hence, district-level unemployment is influenced positively by unemployment in neighboring districts. The spatial autocorrelation coefficient indicates the impact of regional effects that affect a region consisting of more than one district. Examples in the context of regional unemployment are exogenous shocks as the closure of a production site. The spatial autoregressive coefficient of the dynamic model is also significant and positive. The same holds for the pure time-dynamic effect. This result underlines the fact that our data exhibit not only spatial but also temporal autocorrelation. Contrary to this, the combined spatial-time effect is negative and significant.

Furthermore, the results are fairly sensitive to the choice of the spatial weights matrix. In the spatial econometric literature, Bell and Bockstael (2000) (among others) find that estimation results are more sensitive to the specification of the spatial weights matrix than to the estimation technique. According to information criteria, the binary spatial weights matrix captures the spatial structure of the data in the best way for the static model while the distance decay function is most appropriate in case of the dynamic model.

2.5.2 Differences between East and West Germany

Due to German history, it is worthwhile to analyze the differences between the Western and Eastern part of the country. We use a two-regime regression, i.e. we estimate the model for both parts separately. This procedure rests on the assumption that coefficients of the explanatory variables differ between East and West Germany. From an economic perspective, we find no reason why a particular coefficient, for example the coefficient of the employment share in manufacturing, should be similar in all East German and all West German districts. We also tested for the coefficients to be different between East and West Germany. The test results show that most of the coefficients differ significantly between East and West Germany.

The Eastern part of Germany consists of 87 districts and the Western part consists of 325 districts. Note that the spatial model specifications of the separate regressions

dependent variable: u_{it}				
	basic	static		
		binary	distance ($\tau = 0.02$)	Molho (1995) ($\eta = 0.01$)
EG_{it}	-0.066^{***}	-0.033^{***}	-0.04^{***}	-0.05^{***}
	(-7.12)	(-6.2)	(-5.41)	(-6.15)
$\%IND_{it}$	-0.11^{***}	-0.071^{***}	-0.08^{***}	-0.09^{***}
	(-7.95)	(-7.35)	(-7.11)	(-7.05)
$\%ENERW_{it}$	0.17^{***}	0.098^{**}	0.08	0.12^{*}
	(2.6)	(1.98)	(1.47)	(1.93)
$\%CON_{it}$	-0.29^{***}	-0.133^{***}	-0.12^{***}	-0.17^{***}
	(-11.85)	(-10.73)	(-5.58)	(-7.46)
$\%HOT_{it}$	0.16^{***}	0.072^{*}	-0.01	0.09^{*}
	(2.96)	(1.95)	(-0.17)	(1.96)
$\%FIN_{it}$	0.17^{***}	0.046	0.102^{**}	0.14^{***}
	(3.06)	(1.13)	(2.21)	(2.75)
$\%PUB_{it}$	0.12^{***}	0.053^{***}	0.056^{**}	0.073^{***}
	(4.36)	(2.74)	(2.49)	(3.05)
$YOUNG_{it}$	0.35^{***}	0.021	-0.008	0.057
	(9.75)	(0.96)	(-0.24)	(1.64)
OLD_{it}	-0.16^{***}	-0.13^{***}	-0.2^{***}	-0.22^{***}
	(-5.86)	(-7.28)	(-8.03)	(-8.73)
$H0_{it}$	0.103^{***}	0.098^{***}	0.088^{***}	0.089^{***}
	(3.8)	(7.78)	(4.52)	(4.15)
$H1_{it}$	0.081^{***}	0.081^{***}	0.079^{***}	0.084^{***}
	(4.14)	(7.56)	(4.72)	(4.74)
REG_{it}	0.17^{***}	0.08^{***}	0.14^{***}	0.11^{***}
	(4.44)	(3.35)	(4.44)	(3.96)
$DEBTR_{it}$	0.054^{**}	0.015	0.02	0.026
	(2.2)	(0.87)	(0.97)	(1.16)
λ	—	0.83^{***}	0.79^{***}	0.78^{***}
		(71.59)	(16.41)	(14.56)
ρ	—	-0.46^{***}	0.67^{***}	0.71^{***}
		(-13.68)	(8.77)	(9.06)
σ^2	0.61	0.34	0.44	0.5
log-like	-4123.08	-3274.95	-3361.05	-3525.43
AIC	2.23	1.78	1.82	1.82
BIC	2.25	1.80	1.85	1.85
obs.	3708	3708	3708	3708

Notes: *t*-statistics are in parentheses. t-statistics for the static model are computed according to Anselin (1988). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 2.6 Regression results of regional unemployment model - basic and static model specification for the period from 1999 until 2007

dependent variable: u_{it}			
	binary	distance ($\tau = 0.02$)	Molho (1995) ($\eta = 0.01$)
EG_{it}	-0.055^{***}	-0.063^{***}	-0.069^{***}
	(-7.93)	(-8.98)	(-9.23)
$\%IND_{it}$	-0.021^{**}	-0.011	-0.015
	(-1.99)	(-1.05)	(-1.31)
$\%ENERW_{it}$	0.209^{***}	0.23^{***}	0.24^{***}
	(4.25)	(4.63)	(4.56)
$\%CON_{it}$	0.058^{***}	0.019	0.047^{**}
	(2.7)	(0.87)	(2.03)
$\%HOT_{it}$	-0.08^{*}	-0.02	-0.056
	(-1.85)	(-0.36)	(-1.22)
$\%FIN_{it}$	0.055	0.07	0.065
	(1.16)	(1.39)	(1.29)
$\%PUB_{it}$	0.0107	0.0095	0.0078
	(0.53)	(0.47)	(0.36)
$YOUNG_{it}$	0.046^{*}	-0.0001	0.078^{***}
	(1.74)	(-0.0039)	(2.7)
OLD_{it}	-0.08^{***}	-0.104^{***}	-0.104^{***}
	(-3.8)	(-4.9)	(-4.61)
$H0_{it}$	0.006	0.0091	-0.016
	(0.28)	(0.43)	(-0.71)
$H1_{it}$	-0.0149	-0.014	-0.029^{*}
	(-0.97)	(-0.91)	(-1.73)
REG_{it}	0.0073	0.013	0.013
	(0.28)	(0.51)	(0.48)
$DEBTR_{it}$	0.0088	0.0043	0.0081
	(0.48)	(0.23)	(0.41)
λ	0.5^{***}	0.88^{***}	0.79^{***}
	(26.55)	(42.41)	(32.87)
γ	0.78^{***}	0.78^{***}	0.8^{***}
	(49.04)	(52.11)	(55.29)
δ	-0.42^{***}	-0.68^{***}	-0.71^{***}
	(-15.98)	(-17.55)	(-14.31)
σ^2	0.27	0.27	0.31
log-like	-2270.7	-1251.6	-1444.4
AIC	1.39	0.77	0.89
BIC	1.42	0.80	0.92
obs.	3296	3296	3296

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010a). λ is the spatial autoregressive coefficient, γ captures the pure time effect and δ captures the combined spatial-time effect. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively. The reduced number of observations results from Lee and Yu's transformation approach.

Table 2.7 Regression results of regional unemployment model - dynamic model specification for the period from 1999 until 2007 take only the spatial relations within the areas into account, but not the spatial interactions between them. The results for East Germany are in Table 2.8 while the results for Western Germany can be found in Table 2.9.

Employment growth and the employment share in manufacturing are negatively related to regional unemployment in both parts of Germany. Beyond that, the estimation results differ between East and West Germany. Firstly, the employment share of the construction industry has only a significant (negative) impact on regional unemployment rates in Eastern Germany. During the 1990s, economic growth in Eastern Germany was strongly driven by an expansion of the construction industry. Though, it contracted towards the end of the decade (Davies and Hallet (2001)). In 2000, the share of the construction industry in gross value added (using current prices) amounted to 8.1% in East Germany (including Berlin) whereas it amounted to only 4.7% in West Germany.[13] Secondly, regional unemployment in East Germany is only influenced by some of the factors which we account for in our model. Only the educational variables, the number of business registrations and the employment share in hotels and restaurants have a significant (positive) impact on Eastern German regional unemployment rates. Contrary to this, the age variables as well as the employment shares in financial business and in public administration are significant in our model for Western Germany. Thirdly, the influence of the employment share in hotels and restaurants is positive in East Germany while it is negative in West Germany.

In line with our results for the whole country, we find that both spatial and temporal dynamics have to be accounted for when modeling regional unemployment. According to information criteria, the binary contiguity matrix captures the spatial relations in the best way for East Germany. In contrast, the binary contiguity matrix captures the spatial structure best for the static model while the Molho (1995) matrix is the best weights matrix in case of the dynamic model for Western Germany. The spatial coefficients are mostly significant for both parts of Germany. The signs of the coefficients are in line with the results for the whole country. Both the spatial autoregressive and the spatial autocorrelation coefficient are positive. Likewise, the pure time-dynamic is significant and positive while the combined space-time effect is negative and mostly significant.

2.6 Conclusion

In this paper, we analyze the determinants for regional unemployment in Germany. Regional unemployment rates in Germany are characterized by substantial regional differences. We show that there are significant spatial spillovers in regional unemployment data. To avoid biased and inefficient estimates, we apply a spatial panel model to our data. In addition to spatial dependence, we also control for temporal dynamics in the data by specifying a dynamic model. Our analysis covers both the whole of Germany and East and West Germany separately.

[13] The values are taken from *Volkswirtschaftliche Gesamtrechnungen der Länder*, `http://www.vgrdl.de/Arbeitskreis_VGR/`.

	dependent variable: u_{it}				
	basic	static		dynamic	
		binary	Molho (1995) ($\eta = 0.01$)	binary	Molho (1995) ($\eta = 0.01$)
EG_{it}	-0.084^{***}	-0.057^{***}	-0.078^{***}	-0.089^{***}	-0.1^{***}
	(-4.25)	(-3.6)	(-4.3)	(-5.71)	(-6.08)
$\%IND_{it}$	-0.075^{**}	-0.061^{**}	-0.048	0.0074	0.035
	(-2.09)	(-1.96)	(-1.35)	(0.25)	(1.11)
$\%ENERW_{it}$	0.14	-0.12	0.0059	0.21	0.23
	(0.7)	(-0.68)	(0.03)	(1.23)	(1.3)
$\%CON_{it}$	-0.24^{***}	-0.23^{***}	-0.21^{***}	-0.0094	-0.0061
	(-4.11)	(-6.39)	(-4.63)	(-0.19)	(-0.12)
$\%HOT_{it}$	0.27^{***}	0.13^{**}	0.15^{**}	-0.021	0.0035
	(3.64)	(1.97)	(2.01)	(-0.32)	(0.05)
$\%FIN_{it}$	-0.098	-0.35^{*}	-0.23	-0.26	-0.27
	(-0.45)	(-1.88)	(-1.09)	(-1.59)	(-1.55)
$\%PUB_{it}$	0.066	0.0016	0.029	-0.0024	-0.0059
	(1.58)	(-0.04)	(0.71)	(-0.08)	(-0.18)
$YOUNG_{it}$	0.036	0.12^{*}	0.045	-0.018	-0.015
	(0.4)	(1.72)	(0.53)	(-0.25)	(-0.2)
OLD_{it}	-0.046	-0.08^{*}	-0.069	-0.15^{***}	-0.16^{***}
	(-0.83)	(-1.72)	(-1.31)	(-3.42)	(-3.54)
$H0_{it}$	0.19^{**}	0.14^{**}	0.18^{**}	0.056	0.048
	(2.35)	(2.22)	(2.36)	(0.89)	(0.73)
$H1_{it}$	0.092^{*}	0.069^{*}	0.12^{***}	-0.014	-0.015
	(1.86)	(1.78)	(2.64)	(-0.37)	(-0.38)
REG_{it}	0.086	0.083^{*}	0.13^{***}	-0.029	-0.026
	(1.56)	(1.82)	(2.89)	(-0.74)	(-0.62)
$DEBTR_{it}$	0.11^{**}	0.072	0.06	0.039	0.045
	(2.01)	(1.49)	(1.09)	(0.89)	(0.96)
λ	—	0.67^{***}	0.76^{***}	0.38^{***}	0.33^{***}
		(13.84)	(12.44)	(8.83)	(3.74)
ρ	—	-0.03	0.42^{***}	—	—
		(-0.31)	(2.73)		
γ	—	—	—	0.74^{***}	0.72^{***}
				(21.24)	(23.44)
δ	—	—	—	-0.31^{***}	-0.15
				(-5.29)	(-1)
σ^2	0.79	0.59	0.75	0.37	0.41
log-like	-965.54	-824.36	-868.21	-594.99	-766.51
AIC	2.5	2.15	2.26	1.76	2.25
BIC	2.59	2.24	2.35	1.87	2.36
obs.	783	783	783	696	696

Notes: t-statistics are in parentheses. t-statistics for the static model are computed according to Anselin (1988). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 2.8 Regression results of East German regional unemployment model for period from 1999 until 2007

dependent variable: u_{it}					
	basic	static		dynamic	
		binary	Molho (1995) ($\eta = 0.01$)	binary	Molho (1995) ($\eta = 0.01$)
EG_{it}	-0.043^{***}	-0.016^{**}	-0.036^{***}	-0.06^{***}	-0.077^{***}
	(-4.28)	(-2.26)	(-4.05)	(-7.55)	(-9.34)
$\%IND_{it}$	-0.12^{***}	-0.051^{***}	-0.1^{***}	-0.012	-0.0082
	(-8.9)	(-5.37)	(-8.26)	(-1.12)	(-0.73)
$\%ENERW_{it}$	0.12^{*}	0.047	0.12^{**}	0.22^{***}	0.25^{***}
	(1.91)	(1.11)	(2.11)	(4.51)	(4.9)
$\%CON_{it}$	-0.061^{*}	0.0066	-0.052^{*}	0.07^{**}	0.062^{**}
	(-1.89)	(0.29)	(-1.82)	(2.5)	(2.11)
$\%HOT_{it}$	-0.28^{***}	-0.14^{**}	-0.27^{***}	-0.064	-0.063
	(-3.22)	(-2.28)	(-3.42)	(-0.92)	(-0.87)
$\%FIN_{it}$	0.16^{***}	0.07^{**}	0.13^{***}	0.09^{*}	0.1^{**}
	(3.29)	(2.14)	(2.86)	(1.92)	(2.06)
$\%PUB_{it}$	0.14^{***}	0.08^{***}	0.12^{***}	0.021	0.0062
	(4.22)	(3.46)	(3.83)	(0.69)	(0.2)
$YOUNG_{it}$	-0.04	-0.12^{***}	-0.11^{**}	0.072	0.07
	(-0.73)	(-2.72)	(-2.38)	(1.62)	(1.51)
OLD_{it}	-0.31^{***}	-0.25^{***}	-0.36^{***}	-0.01	-0.03
	(-10.12)	(-8.66)	(-11.79)	(-0.4)	(-1.12)
$H0_{it}$	0.034	0.065^{***}	0.037^{*}	0.014	-0.0002
	(1.19)	(3.31)	(1.72)	(0.6)	(-0.01)
$H1_{it}$	0.038^{*}	0.078^{***}	0.081^{***}	-0.034^{**}	-0.054^{***}
	(1.88)	(4.51)	(4.54)	(-2)	(-3.05)
REG_{it}	0.25^{***}	0.08^{**}	0.2^{***}	0.05	0.056
	(4.65)	(1.98)	(4.37)	(1.25)	(1.35)
$DEBTR_{it}$	0.01	-0.0019	0.0084	-0.0043	-0.012
	(0.41)	(-0.1)	(0.36)	(-0.21)	(-0.58)
λ	—	-0.62^{***}	0.76^{***}	0.47^{***}	0.83^{***}
		(-20.09)	(16.52)	(21.1)	(33.95)
ρ	—	0.96^{***}	0.6^{***}	—	—
		(189.74)	(6.72)	—	—
γ	—	—	—	0.805^{***}	0.84^{***}
				(44.66)	(50.95)
δ	—	—	—	-0.37^{***}	-0.84^{***}
				(-11.5)	(-16.2)
σ^2	0.45	0.26	0.39	0.23	0.25
log-like	-2820.29	-2404.88	-2457.77	-2028.3	-841.5
AIC	1.94	1.66	1.69	1.57	0.66
BIC	1.97	1.69	1.72	1.61	0.7
obs.	2925	2925	2925	2600	2600

Notes: t-statistics are in parentheses. t-statistics for the static model are computed according to Anselin (1988). λ is the spatial autoregressive coefficient and ρ is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 2.9 Regression results of West German regional unemployment model for period from 1999 until 2007

Our results clearly show that the spatial panel model fits our data better than the basic model. Moreover, the dynamic modeling is more appropriate for our analysis of regional unemployment than the static one. Hence, the spatial dynamic panel model is the best model for the analysis of regional unemployment.

Our study leads to several conclusions for policy. Firstly, policy measures to reduce unemployment should account for regional differences in unemployment in Germany. In addition to framework conditions that are set on the national level, policy makers have to take regional differences in economic performance into account. Secondly, according to our results, regional unemployment in Germany is of disequilibrium nature. Hence, our results provide a justification for policy makers to intervene on regional labor markets. However, we find significant spatial spillovers, i.e. political decisions do not only affect the district on which they are targeted but also the neighboring districts. This aspect motivates political cooperation between different districts. The definition of labor market regions as, for example, proposed by Eckey et al. (2007) can offer guidance for this process.

Chapter 3

Spatial weights for labor market applications

This chapter is based on **?**.

3.1 Introduction

Labor market activity is correlated over space as firms do not restrict their recruiting activities to their resident location and job searchers might accept a job in a different area. The spatial correlation of labor market data needs to be taken into account when modeling labor market activity as neglecting spatial dependence yields biased and/or inefficient estimates (see, for example, Anselin and Bera (1998) or Franzese and Hays (2007)). The spatial econometric literature provides an extensive toolkit for estimating and testing spatial models. To incorporate the spatial structure of the data into the model, a spatial weights matrix is defined. Its definition is crucial in spatial econometric models as the results heavily depend on the spatial weights (see, for example, Florax and Rey (1995), Griffith (1996), Stakhivych and Bijmolt (2008) or Páez et al. (2008)).

In this chapter, we investigate which factors drive spatial dependence between regional labor markets. Hence, our aim is to get further insights into the construction of spatial weights in labor market applications. In addition to geographic distance, we consider the sectoral structure of regions, differences in living standard and sociocultural differences as dimensions that explain spatial interaction between labor markets. Geographical distance has a frictional effect on labor market activity and it definitely matters for the decisions of job searchers and firms. However, we argue that geographic distance does not capture the interregional activities of job searchers and firms sufficiently as the fact that two regions are geographically close to each other does not necessarily imply that job searchers find a job in the contiguous region.

Spatial econometric models have been applied in different empirical labor market studies as, for example, Burgess and Profit (2001), Fahr and Sunde (2006), Lottmann (2012a, 2012b), Molho (1995), Aragon et al. (2003), Cracolici et al. (2007) and Longhi et al. (2006). However, most of these contributions specify the spatial weights matrix using geographic information only. To our best knowledge, the sole exceptions are Molho (1995) who specifies the spatial weights matrix using a combination of size of regional labor markets (measured in employment) with a (geographic) distance decay effect. Möller and Aldashev (2007) as well as Lottmann (2012b) utilize commuter streams to measure spatial relations between regional labor markets. Our analysis contributes to the existing literature by empirically investigating the issue of defining spatial weights for labor market applications. We propose factors driving spatial dependence in regional labor markets. Hence, our analysis gives an indication for determinants of both the commuting and migration decisions of workers and relocation decisions of firms.

We base the study on the matching model which is a standard macroeconomic tool to analyze the dynamics on labor markets. The matching function describes the relationship between the flow of matches and both the stock of unemployed and the stock of vacancies. Burda and Profit (1996) introduce spatial interactions into the matching function where matches do not only depend on the local labor market situation but also on that in neighboring regions. Lottmann (2012b) empirically shows that there is significant spatial autocorrelation in German matching data. We combine matching data with a huge regional data set providing information on 412 German districts. Our data set covers the years 2008 and 2009. The matching data is quarterly data while the regional data base contains yearly information.

In the applied spatial econometric literature, the spatial lag model and the spatial error model are widely used. They are mostly specified using one type of spatial weights matrix only. As we want to take more than one transmission channel of spatial dependence into account, we apply a higher-order spatial autoregressive model with spatially autocorrelated errors to our data. This model allows to include different spatial weights matrices, based on alternative concepts of distance, to construct various spatially lagged terms. Badinger and Egger (2011) propose a two-step generalized moments estimation framework for the higher-order spatial autoregressive model with spatially autocorrelated errors allowing for heteroscedastic errors which we adopt here.

Our results show that geographic distance alone is not able to sufficiently capture the spatial structure in the job creation process. But economic distance based on both the sectoral structure and the living standard entails additional power in explaining spatial interactions between regional labor markets.

The remainder of the chapter is organized as follows: The second section introduces the theoretical context of spatial dependence in matching functions. The data is described in the third section while the fourth section explains how we define economic distance on labor markets and how we compute spatial weights based on these distances. The econometric model is presented in section five and the sixth section is dedicated to the empirical results. Finally, the last section concludes.

3.2 Spatial dependence in matching functions

To analyze which factors determine spatial dependence between regional labor markets, we need a labor market model that is able to capture the regional dimension of labor market activities. We use the matching function which is a standard tool for the analysis of labor market flows. The labor market is assumed to be a decentralized market where it takes time and resources for jobless individuals and vacant positions to find each other. This two-sided search process can be summarized by a well-behaved function in which the number of new hires (matches) is determined by the stock of unemployed individuals and the stock of vacancies, i.e.

$$\boldsymbol{m} = f(\boldsymbol{U}, \boldsymbol{V}) \tag{3.1}$$

where $\boldsymbol{m}$ denotes the matches, $\boldsymbol{U}$ and $\boldsymbol{V}$ represent the unemployment stock and the vacancy stock, respectively.[1] The search and matching process is characterized by trading frictions, imperfect information and heterogeneities. As Fahr and Sunde (2006) point out, geographical distance can be interpreted as natural friction to search and, therefore, it has frictional effects on the job creation process.

The standard matching function does not control for interregional linkages in the matching process. The empirical matching literature starts with the definition of matching functions on an aggregate level (Blanchard and Diamond (1989), Van Ours (1991), Burda and Wyplosz (1994), Berman (1997)). This approach implicitly assumes that there is one aggregate labor market, i.e. every job searcher is expected to accept a suitable job in the whole country. As this assumption is not generally true, authors started to estimate matching functions on a regional level (Burda (1993), Coles and Smith (1996), Anderson and Burgess (2000), Fahr and Sunde (2006a, 2006b, 2009)). But these contributions treat regional labor markets as isolated markets and do not take into account any linkages between regions. However, labor market activities are not necessarily limited to one specific region as both job searchers and firms do not restrict their search and recruiting activities to their resident location, but search and recruit interregionally. They pursue these activities as long as the expected gains from search exceed search costs.

Though, a few authors propose extensions of the matching function to control for the spatial dimension of the job creation process. To our best knowledge, Burda and Profit (1996) are the first who derive, based on a model of nonsequential search with endogenous search intensity, a matching function relating matches not only to local labor market conditions but also to those in neighboring regions. Burgess and Profit (2001) extend this work and provide empirical evidence on the nature of spatial externalities in the matching process in Great Britain. Their results show significant spatial dependence between local labor markets that decays with distance between them. They explain their finding with the underlying assumption that search costs increase with distance. In order to study regional mobility induced by regional disparities in labor market conditions, Fahr and Sunde (2006) propose an extension of the matching

[1] We use capital letters to denote the stocks and small letters for the flow variables.

framework allowing to study job competition both across regions and between different states of activity. Furthermore, Manning and Petrongolo (2011) propose another extension of the matching model based on a model of optimizing search behavior across space. The aim of their study is to answer the question how 'local' labor markets are. To get insights into this issue, they treat space as continuous instead of consisting of a finite number of non-overlapping areas. Their results suggest that the cost of distance is relatively high as, for example, the utility of being offered a job decays at an exponential rate around 0.3 with distance (in kilometers) to the job.

When spatial dependence is incorporated into empirical models, a spatial weights matrix has to be defined. In the applied spatial econometric literature, the spatial weights matrix is mostly specified using geographic information in the data. We argue that geographic distance is not the only factor driving spatial dependence in the job creation process. The fact that two regions are geographically close to each other does not necessarily imply that job searchers find a job in the contiguous region. But as the decision of a job searcher to work in another region involves commuting or migration, this decision is influenced by personal and socioeconomic factors. Therefore, we also consider dimensions of economic distance for the definition of spatial weights. A proper definition of the spatial weights matrix is essential as the parameter estimates and statistical inference heavily depend on the definition of this matrix (see, for example, Florax and Rey (1995), Griffith (1996), Stakhivych and Bijmolt (2008) and Páez et al. (2008)).

3.3 Data

We need two different kinds of data for the analysis. On the one hand, we use data on matches, the unemployment stock and the stock of vacancies to estimate the matching function. This data is provided by the Federal Employment Office (*Bundesagentur für Arbeit*). On the other hand, we utilize the regional data base of the Federal Statistical Office (*Statistisches Bundesamt*) for the definition of different dimensions of economic distance. Based on these distances, we compute the entries of the spatial weights matrix.

The matching data set contains quarterly information while the data for the definition of economic distances is yearly data. The data is available for all 412 districts of Germany (*Kreise und kreisfreie Städte*) which correspond to the German NUTS III regions.[2] The data set covers only the years 2008 and 2009 for two reasons. Firstly, the

[2]NUTS (French abbreviation) stands for "Nomenclature of Territorial Units for Statistics", and is a hierarchical classification of region units for statistical purposes.

matching data is not completely available for earlier years.[3] Secondly, the data for the different dimensions of economic distance is only available until 2009. As the estimation of the higher-order spatial autoregressive model is computationally demanding, we restrict the empirical analysis to 2009. However, we use the data of 2008 for robustness checks. Table C.1 in Appendix C shows the summary statistics of the labor market variables.

The underlying definition of unemployment is based on regulations in German Social Security Code (*Sozialgesetzbuch*) because a person who is officially considered as unemployed is entitled to social benefits. In Germany, firms are not obliged to report their vacant positions to the Federal Employment Office. Therefore, the registered vacancies represent only a fraction of the aggregate supply of vacant positions. We use the "outflows from unemployment into gainful employment" as measure for the matches. The "outflows from unemployment" in general also include individuals entering into part-time employment, into labor-market policy measures, or people leaving the labor force which we do not consider as successful matches.[4]

3.4 Definition of economic distance and spatial weights

3.4.1 General concepts for the construction of spatial weights

The spatial weights matrix is a nonstochastic matrix that defines exogenously the spatial structure of observations. It is standard in the spatial econometric literature to assume that the elements on the main diagonal are zero because a region cannot be a direct neighbor of itself. Note that we need row-normalized spatial weights matrices for the estimation of the higher-order spatial model. Hence, we have at least to standardize the distances such that the row entries sum up to one.

Row-standardization of spatial weights is commonly used in the applied spatial econometric literature. Operations with row-standardized spatial weights matrices can be interpreted as an averaging of neighboring values (Anselin and Bera (1998)). Though, a row-standardized spatial weights matrix implies that every region is subject to the same total amount of influence from all other regions. In addition to that, the influence of region j on region i decreases with the number of regions influencing region i (Leenders (2002)). Elhorst (2010) discusses the standardization with the largest

[3]This is due to one of the most important labor market reforms ("*Hartz* reform") whose last part became effective at the beginning of 2005. This reform merged unemployment assistance and social welfare to a new form of income support and concurrently changed administrative responsibility. The change in administrative responsibility entails that not all responsible regions reported their data to the Federal Employment Office.

[4]In the empirical matching literature, alternative measures for the matches as, for example, the number of newly filled vacancies are used. Broersma and Van Ours (1999) provide a detailed discussion of different matching measures.

characteristic root of the spatial weights matrix as an alternative to the standard row-normalization. According to Elhorst (2010), the advantage of this normalization lies in the fact that the mutual proportions between the elements of the spatial weights matrix remain unchanged. We experimented with this alternative normalization method and got qualitatively the same results.

There are two different concepts of deriving spatial weights using some distance measure which are commonly used in the applied spatial econometric literature. On the one hand, a distance-decay function can be applied and, on the other hand, a cut-off value concept is possible. A distance decay function has to be a monotonically decreasing function in the argument and, in addition to that, positive arguments have to be strictly associated with nonnegative values (Klotz (2004)). Examples for distance decay functions ($DDF(d_{ij})$) are $DDF(d_{ij}) = 1/d_{ij}$, $DDF(d_{ij}) = 1/d_{ij}^2$ or $DDF(d_{ij}) = exp(-\eta d_{ij})$ where d_{ij} is some distance between regions i and j. The decay parameter η of the exponential distance decay function is the rate by which the impact of a neighbor attenuates with distance (Klotz (2004)).

To apply the cutoff value concept, we have to define a cutoff value d^*. Then, the spatial weights are defined by

$$w_{ij} = \begin{cases} 0; & d_{ij} > d^* \\ DDF(d_{ij}); & d_{ij} < d^*. \end{cases} \tag{3.2}$$

Note that, depending on the empirical context, the weights for $d_{ij} < d^*$ can also be equal to zero while the weights for $d_{ij} > d^*$ can equal the value of the distance decay function.

3.4.2 Geographic and economic distance

Before we construct the entries of the spatial weights matrix, we define different dimensions of distance. The most obvious dimension is geographic distance. Geographic distance clearly affects search costs and, hence, the search activities of job searchers. The applied spatial econometric literature uses different types of geographic distance as, for example, simple contiguity and great circle distances. However, spatial dependence in labor market activity is not only influenced by geographic distance but also by distance in economic terms. Therefore, we consider three dimensions of economic distance in our analysis.

Firstly, we utilize the sectoral decomposition of districts to define economic distances. The sectoral composition of a region plays a role especially in job search. As an example, think of a mining worker who gets unemployed and assume that he wants to stay in his profession. We expect him to look for a new job in his region of residence in the first place. In case he does not find a suitable position, he looks for a job in other regions. However, he will only take those regions into account that offer appropriate jobs for his qualification and his work experience. To measure the sectoral distance, we use data on the share of employed individuals (subject to social security contributions) in every sector. A list of the sectors whose definition corresponds to the NACE classification of economic activities is provided in Appendix C.[5]

Secondly, we consider regional differences in living standards. We consider both disposable income of private households per capita and per capita gross regional product (GRP) to measure the standard of life. The standard deviation of the disposable income amounts to 2,392 euro while it is 10,195 euro in case of the per capita gross regional product in 2009.[6] Hence, there are substantial differences in the level of welfare between German districts. Unfortunately, there is no available data on prices for German districts in order to express income in real terms.

Thirdly, we take sociocultural aspects into account as they are expected to affect the job searcher's decision to migrate for a new job. Peri (2004) shows that sociocultural variables have an impact on the economic development of regions. Sociocultural differences are determined by many factors as, for example, political orientation, linguistic differences, crime as well as individual perceptions of these issues. However, we are limited to those factors that can be measured and for which regional data is available. In order to proxy for sociocultural differences between regions, we utilize data on the results of the *Bundestag* (national German parliament) elections. The elections for the *Bundestag* are the most important elections in Germany since the *Bundestag* elects the German Chancellor. The election which is relevant for the analysis was in 2009. We use data on the percentage of second votes of the five parties that are represented in the *Bundestag*.

Using the different dimensions of economic distance, we define distance matrices. Assume we have L characteristics for n regions. In our case, $L = 13$ for economic distance with respect to the sectoral decomposition, $L = 1$ for economic distance in terms of living standard (gross regional product per capita or disposable household income per capita) and $L = 5$ for the sociocultural distance. Based on the data, we define a matrix Q which is of dimension $(n \times L)$, i.e.

$$Q = (q_{ij}) \quad \text{for} \quad i = 1, \ldots, n; \quad j = 1, \ldots, L. \tag{3.3}$$

We follow the approach proposed by Parent and LeSage (2008) in order to define the distance matrices. In their contribution on knowledge spillovers between European regions, they apply the Jaffe index (Jaffe (1986)) to measure technological proximity. It is defined by

$$S_{ij} = \frac{\sum_{k=1}^{L} q_{ik} q_{jk}}{\sqrt{\sum_{k=1}^{L} q_{ik}^2 \sum_{k=1}^{L} q_{jk}^2}}. \tag{3.4}$$

[5]NACE is a French abbreviation and stands for "Nomenclature statistique des activités économiques dans la Communauté européenne" and is a statistical standard for the classification of different industries in the European Union.

[6]The underlying data is provided online in the regional data base of the Federal Statistical Office.

Assume that the different dimensions of economic distance span a L-dimensional space in which each district's characteristics constitute a vector. Correspondingly, S_{ij} can be interpreted as the cosine of the angle between the vectors of region i and j. S_{ij} takes on larger values when two regions exhibit a similar sectoral structure, i.e. the angle between the vectors of the regions is smaller. S_{ij} will be close to zero for regions whose sectoral structure differs strongly from each other, i.e. the vectors of the regions are close to being orthogonal to each other. Note that this measure of economic distance implies symmetric distances between districts. Finally, the entries of the associated distance matrix D are defined as

$$D = (d_{ij}) = S_{ij} \quad \text{for} \quad i, j = 1, \ldots, n. \tag{3.5}$$

However, S_{ij} is always one in case of the living standard dimension as $L = 1$. We argue that individuals observe differences in living standard by comparing their own level with that of others. Therefore, we construct the distance matrix for the living standard dimension using absolute distances. Thus, the entries of the distance matrix D are defined by

$$d_{ij} = |q_j - q_i|. \tag{3.6}$$

3.4.3 Constructing spatial weights for labor market data

Based on geographic and economic distances, we construct the entries of different spatial weights matrices. To derive the geographic weights, we utilize great circle distances between districts and apply the exponential distance decay function. Hence, the geographic weights are defined by

$$w_{ij}^{geo} = exp(-\eta d_{ij}). \tag{3.7}$$

We consider different values for the choice of the distance decay parameter taking into account the results of previous studies using German regional labor market data (Lottmann (2012a)). Additionally, we define a binary contiguity matrix where the entries are equal to one when two regions share a common boarder. Otherwise, the spatial weights are equal to zero.

We construct the sociocultural weights using a cutoff value d^*. We argue that sociocultural differences do not play a role for jobs that are within commuting distance because job searchers do not need to change their place of residence when accepting a job within commuting distance. Hence, a different sociocultural orientation of the target region is not relevant for the job searcher's decision. The range of d^* is from 25km until 100km because, according to data from the Microcensus 2008, about 76 % of commuters travel less than 25km to work (Grau (2009)).[7] If the distance between

[7]The Microcensus is a official representative one percent sample survey of the population and economic activities in Germany where approximately 800,000 persons participate.

two regions is greater than the cutoff value, we apply the methodology of Parent and LeSage (2008) again. Thus, the sociocultural weights are

$$w_{ij}^{social} = \begin{cases} 0 & for \quad d_{ij} < d^* \\ \frac{\sum_{k=1}^{L} q_{ik} q_{jk}}{\sqrt{\sum_{k=1}^{L} q_{ik}^2 \sum_{k=1}^{L} q_{jk}^2}} & for \quad d_{ij} \geq d^*. \end{cases} \tag{3.8}$$

Finally, the sectoral weights are defined by distances in sectoral decomposition between districts according to equation (3.4), i.e.

$$w_{ij}^{sect} = S_{ij}, \quad \text{for} \quad i, j = 1, \ldots, n. \tag{3.9}$$

As we need row-standardized weights for the estimation of the higher-order spatial autoregressive model, we apply this standardization to all types of the spatial weights matrices.

3.5 Econometric model

3.5.1 Model specification

The baseline model for the analysis of factors driving spatial dependence in regional labor markets is the pooled panel matching specification. We consider the evolution over time in order to capture the temporal dynamics of the matching process.[8] It is standard in the empirical matching literature to assume a Cobb-Douglas specification for the matching function. Hence, the underlying matching specification is

$$ln(\mathbf{m}_t) = \beta_1 ln(\mathbf{U}_t) + \beta_2 ln(\mathbf{V}_t) + \boldsymbol{\epsilon}_t; \quad t = 1, \ldots, T, \tag{3.10}$$

where $\mathbf{m}_t = (m_{1t}, \ldots, m_{nt})'$ is the $(n \times 1)$ vector of matches, $\mathbf{U}_t$ and $\mathbf{V}_t$ are the $(n \times 1)$ vectors of the unemployment and vacancy stocks, respectively. Furthermore, the logarithm is applied component-by-component. The error term $\boldsymbol{\epsilon}_t$ is assumed to be homoscedastic and uncorrelated.

In order to account for spatial dependence in the data, a spatial model specification is commonly derived by introducing a spatially lagged error term and/or a spatial lag of the dependent variable into the model. Hence, these spatial models allow to analyze only one transmission channel of spatial interaction. The omission of relevant transmission channels of spatial dependence results in biased estimates (Badinger and Egger (2011)). Therefore, we apply a higher-order spatial autoregressive model with spatially autocorrelated error terms ($SARAR(R, S)$) to the matching function as this model allows us to analyze which transmission channels of spatial dependence are significant for the job creation process. As Badinger and Egger (2012) point out, the specification

[8] Badinger and Egger (2012) propose an estimation methodology for higher-order spatial autoregressive panel data error component models under the assumption of homoscedastic error terms.

of higher-order spatial processes entails non-trivial issues concerning model specification, the admissible parameter space and the interpretation of the estimates. Elhorst et al. (2012) and LeSage and Pace (2011) provide a detailed discussion of these aspects.

The $SARAR(R, S)$ matching model for each period of time t is

$$ln(\mathbf{m}_t) = \beta_1 ln(\mathbf{U}_t) + \beta_2 ln(\mathbf{V}_t) + \sum_{r=1}^{R} \lambda_r \mathbf{W}_r ln(\mathbf{m}_t) + \boldsymbol{\xi}_t \quad \text{with} \quad \boldsymbol{\xi}_t = \sum_{m=1}^{S} \rho_m \boldsymbol{M}_m \boldsymbol{\xi}_t + \boldsymbol{\epsilon}_t; \quad t = 1, \ldots, T, \quad (3.11)$$

where $\mathbf{W}$ and $\mathbf{M}$ are the $(n \times n)$ nonstochastic spatial weights matrices.

As we consider a pooled panel model, we define the actual sample size N as $N = n * T$. Furthermore, let the modified $(N \times N)$ spatial weights matrices be $\boldsymbol{W}^* = (I_T \otimes \boldsymbol{W})$ and $\boldsymbol{M}^* = (I_T \otimes \boldsymbol{M})$ where I_T is a T-dimensional identity matrix. Stacking the model for all time periods $t = 1, \ldots, T$, we obtain

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \sum_{r=1}^{R} \lambda_r \boldsymbol{W}_r^* \boldsymbol{y} + \boldsymbol{\xi} \quad \text{with} \quad \boldsymbol{\xi} = \sum_{m=1}^{S} \rho_m \boldsymbol{M}_m^* \boldsymbol{\xi} + \boldsymbol{\epsilon} \quad (3.12)$$

where $\boldsymbol{y} = [ln(\boldsymbol{m}_1)', \ldots, ln(\boldsymbol{m}_T)']'$ is a $(N \times 1)$ vector and $\boldsymbol{X} = [\boldsymbol{X}_1', \ldots, \boldsymbol{X}_T']'$ with $\boldsymbol{X}_t = (ln(\boldsymbol{U}_t), ln(\boldsymbol{V}_t)), t = 1, \ldots, T$ is a $(N \times 2)$ matrix. Furthermore, $\beta = [\beta_1, \beta_2]'$, $\boldsymbol{\xi} = [\boldsymbol{\xi}_1', \ldots, \boldsymbol{\xi}_T']'$ and $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_1', \ldots, \boldsymbol{\epsilon}_T']'$. Hence, we allow for R different spatial lags of the log matches and a spatial autoregressive process of order S for the errors. We refer to the vector $\overline{\boldsymbol{y}_r} = \boldsymbol{W}_r^* \boldsymbol{y}$ as the r-th spatial lag of the log matches. Equation (3.12) can be rewritten as

$$\boldsymbol{y} = \boldsymbol{Z}\delta + \boldsymbol{\xi} \quad \text{with} \quad \boldsymbol{\xi} = \sum_{m=1}^{S} \rho_m \boldsymbol{M}_m^* \boldsymbol{\xi} + \boldsymbol{\epsilon} \quad (3.13)$$

where the $(N \times (2 + R))$ matrix $\boldsymbol{Z}$ is given by

$$\boldsymbol{Z} := (\boldsymbol{X}, \overline{\boldsymbol{Y}}) \quad (3.14)$$

with $\overline{\boldsymbol{Y}} = (\overline{\boldsymbol{y}_1}, \ldots, \overline{\boldsymbol{y}_R})$ and $\delta = (\beta', \lambda')'$ where λ denotes the $(R \times 1)$ vector of spatial autoregressive parameters, i.e. $\lambda = (\lambda_1, \ldots, \lambda_R)'$.

Besides the $SARAR$ model, the spatial econometric literature also proposes alternative spatial models as, for example, the spatial lag model and the spatial error model. These models can also be extended to more than one spatial lag or spatially autocorrelated error term, respectively. Hence, a higher-order spatial lag model incorporates cross-sectional dependence in the dependent variable through more than one transmission channel of dependence. Likewise, the higher-order spatial error model allows exogenous shocks to affect other regions through different spatial transmission channels. In the spatial econometric literature, the higher-order spatial error model is applied by, among others, Bell and Bockstael (2000) as well as Cohen and Morrison Paul (2007) while the higher-order spatial lag model is applied, for example, by Lacombe (2004). In

type of spatial weights matrix	contiguity	geo ($\eta = 0.01$)	geo ($\eta = 0.02$)	sectoral	income	GRP	social ($d^* = 25$)
$H_0 : \rho = 0$							
p-value	0	0	0	0.297	0	0.029	0.077
$H_0 : \lambda = 0$							
p-value	0.723	3.35e-06	0.001	0	0	0	0
$H_0 : \rho = 0 \quad (\text{allowing} \quad \lambda \neq 0)$							
p-value	0	0	0	0.422	0	0.06	0.127
$H_0 : \lambda = 0 \quad (\text{allowing} \quad \rho \neq 0)$							
p-value	0	0	0	0	0	0	0
$H_0 : \lambda = 0 \& \rho = 0$							
p-value	0	0	0	0	0	0	0

Table 3.1 Test results of specification test by Anselin et al. (1996a)

order to justify the specification of a $SARAR$ model, we apply the specification tests by Anselin et al. (1996a) to the data. We perform these tests using the same spatial weights matrices for both the spatial lag term and the spatial error term. The test results are in Table 3.1. They show that we can clearly reject the hypothesis of no spatial lag dependence for all spatial weights matrices. Furthermore, the test results give an indication for spatial error dependence as we can reject the associated hypothesis in most of the specifications. Finally, the test results provide additional support for the spatial modeling approach as a whole as the hypothesis of the spatial autoregressive and the spatial autocorrelation coefficient jointly being zero is clearly rejected in all specifications.

3.5.2 Estimation of the $SARAR(R, S)$ model

The literature on the estimation of higher-order spatial autoregressive models is, to our best knowledge, fairly sparse. Lee and Liu (2010) propose an efficient generalized methods of moments (GMM) estimation methodology for the $SARAR(R, S)$ model with homoscedastic error terms. Kelejian and Prucha (2010) deal with heteroscedastic errors but for the $SARAR(1, 1)$ model. They propose a two-step generalized moments (GM) estimation framework for this model. Badinger and Egger (2011) generalize the estimation framework of Kelejian and Prucha (2010) to the $SARAR(R, S)$ model with heteroscedastic errors. We apply their estimation methodology to the matching model.

Kelejian and Prucha (2010) and Badinger and Egger (2011) make some assumptions for their estimation methodology. Regarding the spatial weights matrices and the spatial autoregressive parameters the following assumption is imposed (Badinger and Egger (2011)):

Assumption 1. (a) The row and column sums of the spatial weights matrices $\boldsymbol{W}_r^*, r = 1, \ldots, R$ and $\boldsymbol{M}_m^*, m = 1, \ldots, S$ are bounded uniformly in absolute value.[9] (b) The

[9]For more details on the notion of uniform boundedness, see Kelejian and Prucha (2010).

parameters $\lambda_r, r = 1, \ldots, R$ and $\rho_m, m = 1, \ldots, S$ are finite and contained in the admissible parameter space. With row-normalized matrices it holds that $\sum_{r=1}^{R} |\lambda_r| < 1$ and $\sum_{m=1}^{S} |\rho_m| < 1$. This assumption ensures invertibility of $(\boldsymbol{I}_N - \sum_{r=1}^{R} \lambda_r \boldsymbol{W}_r^*)$ and $(\boldsymbol{I}_N - \sum_{s=1}^{S} \rho_s \boldsymbol{M}_s^*)$ and, thus, the dependent variable, i.e. the log matches, and the term $\boldsymbol{\xi}_t$ are uniquely defined.

Moreover, we assume the error terms to be heteroscedastic. Therefore, the error terms are assumed to satisfy the following assumption.

Assumption 2. The error terms are independently distributed with $E(\epsilon_i) = 0$ and $E(\epsilon_i^2) = \sigma_i^2$ for $i = 1, \ldots, N$. Hence, the variance-covariance matrix of $\boldsymbol{\epsilon}$ is given by

$$Cov(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \sigma_i^2 I_N, \quad i = 1, \ldots, N. \tag{3.15}$$

As our results show that the highest dimensional model is a $SARAR(2,2)$ in our application, we present the estimation methodology assuming that $R, S = 2$.[10] The cases of either $R = 1$ or $S = 1$ are contained as special cases.

The estimation procedure of Badinger and Egger (2011) consists of two main steps. In the first step, the model (3.13) is estimated by two-stage least squares (2SLS) using instruments $\mathbf{H}$ ignoring spatial autocorrelation in the error term. Kelejian and Prucha (2010) as well as Badinger and Egger (2011) show that the optimal instruments $\mathbf{H}$ consist of a subset of linearly independent columns of $(\mathbf{X}, \mathbf{WX}, \mathbf{W}^2\mathbf{X}, \ldots, \mathbf{W}^p\mathbf{X})$ where p is a pre-selected finite constant. Kelejian et al. (2004) suggest to set $p \leq 2$. Hence, the 2SLS estimator of δ is defined by

$$\widetilde{\delta} = (\widehat{\boldsymbol{Z}}'\widehat{\boldsymbol{Z}})^{-1}\widehat{\boldsymbol{Z}}'\boldsymbol{y}, \tag{3.16}$$

where

$$\widehat{\boldsymbol{Z}} = \boldsymbol{H}(\boldsymbol{H}'\boldsymbol{H})^{-1}\boldsymbol{H}'\boldsymbol{Z} = \boldsymbol{P}_H\boldsymbol{Z}. \tag{3.17}$$

In the second step, the spatial autocorrelation parameters ρ_1 and ρ_2 are estimated using a GM estimator based on the residuals $\widetilde{\boldsymbol{\xi}} = \boldsymbol{y} - \boldsymbol{Z}\widetilde{\delta}$ from the first step.

Under Assumptions 1 and 2, Badinger and Egger (2011) derive the set of moment conditions. The second-order spatial autoregressive process for the error terms entails additional moment conditions that are associated with spatial weights matrices $\boldsymbol{M}_1^*$ and $\boldsymbol{M}_2^*$ as well as combinations of them. Let us define

$$\overline{\boldsymbol{\epsilon}}_s = \boldsymbol{M}_s^*\boldsymbol{\epsilon}. \tag{3.18}$$

Then, the moment conditions for $s = 1, 2$ and $s' > s$ are

$$M_1^{s,s'} : N^{-1}[E(\overline{\boldsymbol{\epsilon}}_s'\overline{\boldsymbol{\epsilon}}_{s'}) - tr(\boldsymbol{M}_{s'}^*[diag_{i=1}^{N} E(\epsilon_i^2)]\boldsymbol{M}_s^{*\prime})] = 0 \tag{3.19a}$$

$$M_2^{s} : N^{-1}E(\overline{\boldsymbol{\epsilon}}_s'\boldsymbol{\epsilon}) = 0, \tag{3.19b}$$

[10] The formulas for the general case are in Badinger and Egger (2011).

where $tr(\bullet)$ denotes the trace operator. For notational convenience, let $\overline{\boldsymbol{\xi}}_s := \boldsymbol{M}_s^*\boldsymbol{\xi}$, $\overline{\overline{\boldsymbol{\xi}}}_{sm} := \boldsymbol{M}_s^*\boldsymbol{M}_m^*\boldsymbol{\xi} = \boldsymbol{M}_s^*\overline{\boldsymbol{\xi}}_m$. From the model specification (3.12), it follows that

$$\boldsymbol{\epsilon} = \boldsymbol{\xi} - \sum_{m=1}^{2} \rho_m \boldsymbol{M}_m^* \boldsymbol{\xi} = \boldsymbol{\xi} - \sum_{m=1}^{2} \rho_m \overline{\boldsymbol{\xi}}_m \quad \text{and} \tag{3.20a}$$

$$\overline{\boldsymbol{\epsilon}} = \boldsymbol{M}_s^*\boldsymbol{\epsilon} = \boldsymbol{M}_s^*(\boldsymbol{\xi} - \sum_{m=1}^{2} \rho_m \boldsymbol{M}_m^* \boldsymbol{\xi}) = \overline{\boldsymbol{\xi}} - \sum_{m=1}^{2} \rho_m \overline{\overline{\boldsymbol{\xi}}}_{sm}. \tag{3.20b}$$

Substituting (3.20a) and (3.20b) into the moment conditions (3.19) gives a system of five equations in ρ_1 and ρ_2 which is

$$\boldsymbol{\gamma} - \boldsymbol{\Gamma}\boldsymbol{\alpha} = \mathbf{0} \tag{3.21}$$

where $\boldsymbol{\alpha} = (\rho_1, \rho_2, \rho_1^2, \rho_2^2, \rho_1\rho_2)'$. γ is a (5×1) vector where its entries associated with moment conditions $M_1^{s,s'}$ are

$$\gamma_{2(s-1)+1} = N^{-1}E[\overline{\boldsymbol{\xi}}_s'\overline{\boldsymbol{\xi}}_s - tr(\boldsymbol{M}_s^* diag_{i=1}^N(\xi_i^2)\boldsymbol{M}_s^{*\prime})] \quad \text{for} \quad s = 1,2 \tag{3.22a}$$

$$\gamma_5 = N^{-1}E[\overline{\boldsymbol{\xi}}_1'\overline{\boldsymbol{\xi}}_2 - tr(\boldsymbol{M}_2^* diag_{i=1}^N(\xi_i^2)\boldsymbol{M}_1^{*\prime})]. \tag{3.22b}$$

The remaining entries of γ, associated with M_2^s, are

$$\gamma_{2(s-1)+2} = N^{-1}E[\overline{\boldsymbol{\xi}}_s'\boldsymbol{\xi}] \quad \text{for} \quad s = 1,2. \tag{3.23}$$

$\boldsymbol{\Gamma}$ is a (5×5) matrix where the entries associated with $M_1^{s,s'}$ are

$$\gamma_{2(s-1)+1,m} = 2N^{-1}E[\overline{\boldsymbol{\xi}}_s'\overline{\overline{\boldsymbol{\xi}}}_{sm} - tr(\boldsymbol{M}_s^* diag_{i=1}^N(\overline{\xi}_{m,i}\xi_i)\boldsymbol{M}_s^{*\prime})] \quad \text{for} \quad s = 1,2; m = 1,2 \tag{3.24a}$$

$$\gamma_{5,m} = 2N^{-1}E[\overline{\boldsymbol{\xi}}_1'\overline{\overline{\boldsymbol{\xi}}}_{2m} - tr(\boldsymbol{M}_2^* diag_{i=1}^N(\overline{\xi}_{m,i}\xi_i)\boldsymbol{M}_1^{*\prime})] \quad \text{for} \quad m = 1,2 \tag{3.24b}$$

$$\gamma_{2(s-1)+1,m+2} = -N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{sm}'\overline{\overline{\boldsymbol{\xi}}}_{sm} - tr(\boldsymbol{M}_s^* diag_{i=1}^N(\overline{\xi}_{m,i}^2)\boldsymbol{M}_s^{*\prime})] \quad \text{for} \quad s = 1,2; m = 1,2 \tag{3.24c}$$

$$\gamma_{5,m+2} = -N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{1m}'\overline{\overline{\boldsymbol{\xi}}}_{2m} - tr(\boldsymbol{M}_2^* diag_{i=1}^N(\overline{\xi}_{m,i}^2)\boldsymbol{M}_1^{*\prime})] \quad \text{for} \quad m = 1,2 \tag{3.24d}$$

$$\gamma_{2(s-1)+1,5} = -2N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{s1}'\overline{\overline{\boldsymbol{\xi}}}_{s2} - tr(\boldsymbol{M}_s^* diag_{i=1}^N(\overline{\xi}_{1,i}\overline{\xi}_{2,i})\boldsymbol{M}_s^{*\prime})] \quad \text{for} \quad s = 1,2 \tag{3.24e}$$

$$\gamma_{5,5} = -2N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{11}'\overline{\overline{\boldsymbol{\xi}}}_{22} - tr(\boldsymbol{M}_2^* diag_{i=1}^N(\overline{\xi}_{1,i}\overline{\xi}_{2,i})\boldsymbol{M}_1^{*\prime})]. \tag{3.24f}$$

The entries of $\mathbf{\Gamma}$ associated with moment conditions M_2^s are

$$\gamma_{2(s-1)+2,m} = N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{sm}'\boldsymbol{\xi} + \overline{\boldsymbol{\xi}}_s'\overline{\boldsymbol{\xi}}_m] \quad \text{for} \quad s = 1,2; m = 1,2 \tag{3.25a}$$

$$\gamma_{2(s-1)+2,2+m} = -N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{sm}'\overline{\boldsymbol{\xi}}_m] \quad \text{for} \quad s = 1,2; m = 1,2 \tag{3.25b}$$

$$\gamma_{2(s-1)+2,5} = -N^{-1}E[\overline{\overline{\boldsymbol{\xi}}}_{s2}'\overline{\boldsymbol{\xi}}_1 + \overline{\overline{\boldsymbol{\xi}}}_{s1}'\overline{\boldsymbol{\xi}}_2] \quad \text{for} \quad s = 1,2. \tag{3.25c}$$

The empirical analogue of the equation system in equation (3.21) is

$$\widetilde{\gamma} - \widetilde{\mathbf{\Gamma}}\boldsymbol{\alpha} = \boldsymbol{v} \tag{3.26}$$

where the elements of $\widetilde{\gamma}$ and $\widetilde{\mathbf{\Gamma}}$ correspond to the elements of γ and $\mathbf{\Gamma}$ with the expectations operator suppressed and the disturbances $\boldsymbol{\xi}$ replaced by consistent estimates $\widetilde{\boldsymbol{\xi}}$. Badinger and Egger (2011) define the GM estimates of ρ_1 and ρ_2 as a solution to

$$\underset{\rho_1,\rho_2}{argmin}[(\widetilde{\gamma} - \widetilde{\mathbf{\Gamma}}\boldsymbol{\alpha})'\widetilde{\mathbf{\Theta}}(\widetilde{\gamma} - \widetilde{\mathbf{\Gamma}}\boldsymbol{\alpha})] = [\boldsymbol{v}'\widetilde{\mathbf{\Theta}}\boldsymbol{v}] \tag{3.27}$$

where $\boldsymbol{v} = (\widetilde{\gamma} - \widetilde{\mathbf{\Gamma}}\boldsymbol{\alpha})$ can be interpreted as vector of regression residuals based on a weighted nonlinear least squares regression of $\widetilde{\gamma}$ on the columns of $\widetilde{\mathbf{\Gamma}}$ (Badinger and Egger (2011)). $\widetilde{\mathbf{\Theta}}$ is a (5×5) weighting function for which Badinger and Egger (2011) discuss the optimal choice. The corresponding formulas are in Appendix C. Furthermore, Badinger and Egger (2011) prove consistency and asymptotic normality of their GM estimators for ρ_1 and ρ_2.

Finally, to improve efficiency, the estimates for ρ_1 and ρ_2 are used to re-estimate the model in an iterative fashion using (feasible) spatial generalized 2SLS estimation of the regression parameters δ. Additionally, Badinger and Egger (2011) derive the joint variance-covariance matrix of the regression parameters δ and the spatial autoregressive parameters $\rho_m, m = 1, \ldots, S$ which can be used to test the $SARAR(R,S)$ model against the alternative of no spatial lags ($SARAR(0,S)$), no spatial errors ($SARAR(R,0)$) or the non-spatial model.

Regarding the spatial coefficients, some remarks have to be made. It lies in the nature of spatial processes that a change in some region i does not only influence the first-order neighbors but has also an impact on second and higher-order neighbors. We illustrate these effects for the second-order spatial autoregressive model (Elhorst et al. (2012)). The second-order spatial autoregressive process reads as (in general notation)

$$y = \lambda_1 W_1 y + \lambda_2 W_2 y + X\beta + \varepsilon \tag{3.28}$$

where y is a $(n \times 1)$ vector of the dependent variables and X is a $(n \times k)$ matrix of regressors. Furthermore, W_1 and W_2 are $(n \times n)$ spatial weights matrices and ε is a $(n \times 1)$ error term. This model can be rewritten as

$$(I_n - \lambda_1 W_1 - \lambda_2 W_2)y = X\beta + \varepsilon \tag{3.29}$$

where I_n is a n-dimensional identity matrix.
Hence, it follows that

$$y = (I_n - \lambda_1 W_1 - \lambda_2 W_2)^{-1}[X\beta + \varepsilon] \tag{3.30}$$

where Assumption 1 needs to be satisfied to ensure the existence of the inverse matrix.

The inverse matrix can be expanded as infinite series (see Elhorst et al. (2012))

$$(I_n - \lambda_1 W_1 - \lambda_2 W_2)^{-1} = (I_n + \lambda_1 W_1 + \lambda_2 W_2) + \sum_{q=2}^{\infty} (\lambda_1 W_1 + \lambda_2 W_2)^q. \tag{3.31}$$

The first term in brackets $(I_n + \lambda_1 W_1 + \lambda_2 W_2)$ represents the first-order effects and the second term $\sum_{q=2}^{\infty}(\lambda_1 W_1 + \lambda_2 W_2)^q$ describes the second and higher-order effects. Thus, the expansion of the inverse matrix shows that it is not possible to separate the impact of the different spatial weights matrices on the second and higher-order effects of changes in some region i from each other. Therefore, the size of the spatial autoregressive and spatial autocorrelation coefficients does not reveal the amount of influence of the associated spatial weights matrix. Hence, we can only use the spatial coefficient estimates to decide which distance dimension has a significant impact on spatial interactions in the matching process.

3.6 Empirical results

When specifying a higher-order model, the first issue to consider is the question of how many spatial lag terms and how many spatial error terms are appropriate for the empirical model. Therefore, we start with the spatial autoregressive model with spatially autocorrelated errors both of order one, i.e. the $SARAR(1,1)$ model. On the one hand, we estimate the $SARAR(1,1)$ model using the same spatial weights matrices both for the spatial lag term and the spatial error term. Thus, we analyze whether the different dimensions of (economic) distance are able to capture the spatial dependence in matching functions at all. The regression results of these regressions are in Table 3.2. On the other hand, we estimate the $SARAR(1,1)$ model using combinations of geographic weights with economic distance weights for the spatial lag and the spatial error term where the results are in Table 3.3. In a next step, we extend this model to allow for two spatial lags which enables us to investigate the joint effects of different transmission channels of spatial dependence. The associated results are in Table 3.4. They show that the spatial interaction in the job creation process can be captured by two different dimensions of distance. Therefore, our results give a clear indication that one dimension of spatial interaction does not capture the spatial structure in the matching process sufficiently. Hence, geographic distance alone is not enough to explain the interregional activities of job searchers which is in line with our expectations. The fact that a nearby region offers many open positions does not imply that unemployed find a job there.

This seems like a trivial finding but it is not considering the fact that most of the contributions applying spatial econometric methods to labor market contexts specify the spatial weights matrix based on geographic information.

Then, we also extend the $SARAR(1,1)$ model to allow for second-order spatial autoregressive error terms which gives results that are not contained in the admissible parameter space as required by Assumption 1. Thus, there is only one channel of spatial dependence in the error terms. Spatially autocorrelated errors describe exogenous effects which are not idiosyncratic to one specific region but also affect neighboring regions. The definition of neighborhood is determined by the entries of the corresponding spatial weights matrix. According to our results, the spatial dependence in the errors is governed through geographic distance. Therefore, we can conclude that the $SARAR(2,1)$ model is the appropriate higher-order model to capture spatial interactions in the matching process.[11]

Based on the estimation results of the $SARAR(2,1)$ model (Table 3.4) in conjunction with the results of the $SARAR(1,1)$ models (Tables 3.2 and 3.3), we identify which channels of spatial dependence are significant in the matching process. Besides the sociocultural weights using a cutoff distance of 25km, our economic distance weights produce significant spatial coefficients. However, it is open whether this finding is of general truth or whether it simply indicates that our proxy for sociocultural differences is poor. The results of the $SARAR(1,1)$ model with different spatial weights for the spatial lag term and the spatial error term as well as the results of the $SARAR(2,1)$ model confirm the conclusion that both the sectoral structure and the level in living standard entail explanatory power for spatial linkages between regional labor markets.

As expected from matching theory, the estimated matching elasticities on both stock variables are positive. However, the matching elasticities on vacancies are not significant. This might be a specific feature of the data of 2009 as our robustness checks show significant matching elasticities with respect to vacancies in 2008. Interestingly, the matching elasticities are very similar across different model specifications. The estimated unemployment elasticity amounts to 0.77 in case of the $SARAR(1,1)$ model using the same spatial weights matrices for the spatial lag and the spatial error term. This means that an increase of the unemployment stock by 1% results in an increase of matching by 0.77%. The estimated elasticities with respect to vacancies are much smaller than those with respect to unemployed in all specifications. This finding might be related to the underreporting of vacanct jobs to the Federal Employment Office. Furthermore, it can be explained by a high vacancy turnover, i.e. vacant jobs are filled within a month and are not counted in the end-of-month stocks. Likewise, they are not captured by quarterly stocks which are aggregates of monthly stocks. As it is standard in the matching literature, we test for constant returns to scale in the matching function. Though, we reject the null hypothesis of constant returns to scale in all model specifications.

The comparison of our estimated matching elasticities with results from other con-

[11] For the sake of completeness, we also consider the $SARAR(2,2)$ model also gives results that are not contained in the admissible parameter required by Assumption 1.

dependent variable: $ln(\boldsymbol{m}_t)$						
$\boldsymbol{W}_1$	contiguity	sectoral	income	GRP	social ($d^* = 25km$)	geo ($\eta = 0.02$)
$\boldsymbol{M}_1$	contiguity	sectoral	income	GRP	social ($d^* = 25km$)	geo ($\eta = 0.02$)
$ln(\boldsymbol{U}_t)$	0.767^{***}	0.768^{***}	0.769^{***}	0.769^{***}	0.767^{***}	0.783^{***}
	(56.23)	(47.95)	(48.39)	(48.08)	(47.85)	(57.05)
$ln(\boldsymbol{V}_t)$	0.012	0.011	0.011	0.011	0.012	0.013
	(0.83)	(0.70)	(0.68)	(0.67)	(0.72)	(0.99)
λ_1	-0.082^{***}	-0.084^{***}	-0.084^{***}	-0.083^{***}	-0.075	-0.109^{***}
	(-9.21)	(-6.78)	(-6.88)	(-6.75)	(-0.82)	(-9.94)
ρ_1	0.391^{***}	-0.567^{***}	-0.511^{***}	-0.895^{***}	0.989^{***}	0.074^{***}
	(17.39)	(-7.77)	(-3.02)	(-27.3)	(20.33)	(63.72)
obs.	1648	1648	1648	1648	1648	1648

Notes: t-statistics are in parentheses. λ_1 is the spatial autoregressive coefficient and ρ_1 is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 3.2 Estimation results of the SARAR(1,1) matching specification with the same weights for the spatial lags and spatial errors (2009)

tributions in the empirical matching literature has two main difficulties. Firstly, we show in an earlier contribution (Lottmann (2012b)) that neglecting spatial dependencies in matching data yields upward-biased coefficients. Most contributions in the empirical matching literature do not take into account spatial linkages in the data. Therefore, the coefficient estimates are not comparable as we account for the spatial structure in the data. Secondly, the contributions in the literature differ not only with respect to the examined country and time period but also with respect to different data definitions for the unemployment stock, vacancy stock and especially the definition of the matches.

Certainly, the time horizon of the analysis is rather short. Therefore, we perform the same regressions using quarterly data for 2008 as robustness check. The regression results of 2008 are in Tables C.2, C.3 and C.4 in Appendix C. The results of 2008 differ from those of 2009 in some respects. Firstly, the matching elasticities with respect to the vacancy stock are mostly significant in the analysis for 2008. Secondly, the weakness of the sociocultural weights using a cutoff distance of 25km is not supported by the results of 2008. However, the other cutoff values for the sociocultural distance do not produce significant spatial coefficients at all, neither for 2009 nor for 2008.

dependent variable: $ln(\boldsymbol{m}_t)$				
$\boldsymbol{W}_1$	contiguity	contiguity	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)
$\boldsymbol{M}_1$	sectoral	income	sectoral	income
$ln(\boldsymbol{U}_t)$	0.767^{***}	0.774^{***}	0.773^{***}	0.77^{***}
	(56.23)	(47.44)	(47.29)	(47.62)
$ln(\boldsymbol{V}_t)$	0.012	0.003	0.009	0.01
	(0.83)	(0.18)	(0.58)	(0.59)
λ_1	-0.082^{***}	-0.083^{***}	-0.089^{***}	-0.085^{***}
	(-9.21)	(-7.14)	(-7.1)	(-6.84)
ρ_1	0.391^{***}	-0.688^{***}	-0.612^{***}	-0.654^{***}
	(17.39)	(-3.98)	(-7.83)	(-3.76)
obs.	1648	1648	1648	1648

Notes: t-statistics are in parentheses. λ_1 is the spatial autoregressive coefficient and ρ_1 is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively. We tried both the geographic weights for the spatial lag term and the economic distance weights for the spatial error term and vice versa. As the results are qualitatively similar, we only report one combination in this table. The full results can be obtained from the author upon request.

Table 3.3 Estimation results of the SARAR(1,1) matching specification using geographic weights in combination with economic distance weights (2009)

	dependent variable: $ln(\boldsymbol{m}_t)$							
$\boldsymbol{W}_1$	contiguity	contiguity	contiguity	contiguity	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)
$\boldsymbol{W}_2$	sectoral	income	GRP	social ($d^* = 25km$)	sectoral	income	GRP	social ($d^* = 25km$)
$\boldsymbol{M}_1$	contiguity	contiguity	contiguity	contiguity	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)
$ln(\boldsymbol{U}_t)$	0.777^{***}	0.78^{***}	0.778^{***}	0.777^{***}	0.782^{***}	0.782^{***}	0.782^{***}	0.782^{***}
	(53.92)	(54.28)	(53.87)	(53.93)	(56.37)	(57.22)	(56.45)	(56.6)
$ln(\boldsymbol{V}_t)$	0.013	0.014	0.013	0.013	0.011	0.015	0.012	0.012
	(0.91)	(1.02)	(0.9)	(0.94)	(0.8)	(1.1)	(0.87)	(0.88)
λ_1	-0.039^{*}	-0.01	-0.038^{*}	-0.033	-0.333^{***}	0.025	-0.271^{**}	-0.241^{**}
	(-1.81)	(-0.49)	(-1.76)	(-1.53)	(-2.7)	(0.3)	(-2.35)	(-2.05)
λ_2	-0.06^{**}	-0.095^{***}	-0.061^{**}	-0.067^{**}	0.228^{*}	-0.135	0.165	0.135
	(-2.23)	(-3.75)	(-2.3)	(-2.51)	(1.82)	(-1.62)	(1.41)	(1.13)
ρ_1	0.375^{***}	0.363^{***}	0.374^{***}	0.373^{***}	0.764^{***}	0.73^{***}	0.759^{***}	0.756^{***}
	(15.72)	(13.96)	(15.63)	(15.19)	(41.75)	(93.33)	(45.96)	(47.17)
obs.	1648	1648	1648	1648	1648	1648	1648	1648
Wald λ (p-value)	0	0	0	0	0	0	0	0

Notes: t-statistics are in parentheses. λ_1 and λ_2 are the spatial autoregressive coefficients associated with the spatial weights matrices W_1 and W_2, respectively. ρ_1 is the spatial autocorrelation coefficient associated with the spatial weights matrix M_1. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table 3.4 Estimation results of the $SARAR(2,1)$ matching specification using both geographic weights and economic distance weights for the spatial lags in combination with geographic weights for the spatial error term (2009)

To facilitate the interpretation of the results, we would like to apply some criterion for model comparison. The standard criteria as Akaike's information criterion and the Bayesian information criterion cannot be applied in our setting as we do not have a likelihood function. Though, the literature provides similar model selection criteria for models that are estimated using the methods of moments methodology. Andrews and Lu (2001) propose consistent model and moment selection procedures for GMM estimation. The starting point of their methodology is a vector of parameters to be estimated. Then, they define a selection vector for both the selection of moments and the selection of some parameters. Based on these selection vectors, they define their model and moment selection criterion. In order to obtain a model where all parameters are stacked into one vector, we rewrite model (3.12) as

$$\boldsymbol{y} = \boldsymbol{Z}\delta + (I_N - \sum_{s=1}^{S} \rho_s \boldsymbol{M}_s^*)^{-1} \boldsymbol{\epsilon}. \tag{3.32}$$

When applying the expansion of the inverse matrix in equation (3.31) to model (3.32), it is obvious that we cannot stack all parameters into one vector. Therefore, the definition of selection vectors is unfeasible. Furthermore, as the estimation procedure of the $SARAR(R, S)$ consists of two steps, the estimation results of both steps need to be combined into one model selection criterion. We leave this issue for future research.

3.7 Conclusion

In this chapter, we empirically investigate which factors drive spatial dependence between regional labor markets. We get further insights into the definition of spatial weights in labor market applications. The correct definition of spatial weights is indispensable in spatial econometric models as all results are strongly influenced by the spatial weights. We utilize the matching model as it is a very prominent tool to analyze the job creation process. Furthermore, the matching model can be extended to consider the spatial dimension of labor market activity. To account for different dimensions of distance, we apply a higher-order spatial autoregressive model with spatially autocorrelated errors. In addition to geographic distance, we consider different dimensions of economic distance. We construct spatial weights based on the sectoral structure, the living standard as well as sociocultural differences.

Our results clearly show that geographic distance does not capture spatial dependence on labor markets in a sufficient way. The dimensions of economic distance entail additional power in determining spatial correlation in regional labor markets. According to the results, both the sectoral structure and differences in living standard clearly affect spatial interactions in matching functions while sociocultural distance plays a minor role.

Bibliography

ANDERSON, P. M. AND S. M. BURGESS (2000): "Empirical Matching Functions: Estimation and Interpretation Using State-Level Data," *The Review of Economics and Statistics*, 82, 93–102.

ANDREWS, D. W. K. AND B. LU (2001): "Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models," *Journal of Econometrics*, 101, 123–164.

ANSELIN, L. (1988): *Spatial econometrics: Methods and models*, Dordrecht: Kluwer.

ANSELIN, L. AND A. K. BERA (1998): "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics," in *Handbook of applied economic statistics*, ed. by A. Ullah and D. E. A. Giles, New York: Marcel Dekker, 237–289.

ANSELIN, L., A. K. BERA, R. FLORAX, AND M. J. YOON (1996a): "Simple diagnostic tests for spatial dependence," *Regional Science and Urban Economics*, 26, 77–104.

ANSELIN, L., A. K. BERA, R. J. G. M. FLORAX, AND M. J. YOON (1996b): "Simple diagnostic tests for spatial dependence," *Regional Science and Urban Economics*, 26, 77–104.

ARAGON, Y., D. HAUGHTON, J. HAUGHTON, E. LECONTE, E. MALIN, A. RUIZ-GAZEN, AND C. THOMAS-AGNAN (2003): "Explaining the pattern of regional unemployment: The case of the Midi-Pyrénées region," *Papers in Regional Science*, 82, 155–174.

BA (2008): "Qualitätsbericht: Statistik der gemeldeten Stellen," .

BADDELEY, M., R. MARTIN, AND P. TYLER (1998): "European Regional Unemployment Disparities: Convergence or Persistence?" *European Urban and Regional Studies*, 5, 195–215.

BADINGER, H. AND P. EGGER (2011): "Estimation of higher-order spatial autoregressive cross-section models with heteroscedastic disturbances," *Papers in Regional Science*, 90, 213–235.

——— (2012): "Estimation and testing of higher-order spatial autoregressive panel data error component models," *Journal of Geographical Systems*, DOI 10.1007/s10109-012-0174-z.

BALTAGI, B. H. (2005): *Econometric analysis of panel data*, Chichester: Wiley, 3. ed.

——— (2008): *Econometric analysis of panel data*, Chichester and UK: John Wiley & Sons, 4 ed.

BALTAGI, B. H., G. BRESSON, AND A. PIROTTE (2007a): "Panel Unit Root Tests and Spatial Dependence," *Journal of Applied Econometrics*, 22, 339–360.

BALTAGI, B. H. AND L. LIU (2008): "Testing for random effects and spatial lag dependence in panel data models," *Statistics and Probability Letters*, 78, 3304–3306.

BALTAGI, B. H., S. H. SONG, B. C. JUNG, AND W. KOH (2007b): "Testing for serial correlation, spatial autocorrelation and random effects using panel data: Analysis of spatially dependent data," *Journal of Econometrics*, 140, 5–51.

BALTAGI, B. H., S. H. SONG, AND W. KOH (2003): "Testing panel data regression models with spatial error correlation," *Journal of Econometrics*, 117, 123–150.

BELL, K. B. AND N. E. BOCKSTAEL (2000): "Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Microlevel Data," *The Review of Economics and Statistics*, 82, 72–82.

BERMAN, E. (1997): "Help Wanted, Job Needed: Estimates of a Matching Function from Employment Service Data," *Journal of Labor Economics*, 15, S251–S292.

BLANCHARD, O. J. AND P. DIAMOND (1989): "The Beveridge Curve," *Brookings Papers on Economic Activity*, 1989, 1–76.

BREUSCH, T. S. AND A. R. PAGAN (1980): "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics," *The Review of Economic Studies*, 47, 239–253.

BROERSMA, L. AND J. C. VAN OURS (1999): "Job searchers, job matches and the elasticity of matching," *Labour Economics*, 6, 77–93.

BURDA, M. C. (1993): "Labour Markets and Structural Change in Eastern Europe," *Economic Policy*, 8, 102–137.

——— (1994): "Modelling Exits from Unemployment in Eastern Germany: A Matching Function Approach," in *Arbeitsmarktdynamik und Unternehmensentwicklung in Osteuropa: Beiträge eines Workshops des Zentrums für Europäische Wirtschaftsforschung am 4. und 5. März 1993 in Mannheim*, ed. by H. König and V. Steiner, Baden-Baden: Nomos-Verl.-Ges.

BURDA, M. C. AND S. PROFIT (1996): "Matching across space: Evidence on mobility in the Czech Republic," *Labour Economics*, 3, 255–278.

BURDA, M. C. AND C. WYPLOSZ (1994): "Gross worker and job flows in Europe," *European Economic Review*, 38, 1287–1315.

BURGESS, S. AND S. PROFIT (2001): "Externalities in the matching function of workers and firms in Britain," *Labour Economics*, 8, 313–333.

CHOI, I. (2001): "Unit root tests for panel data," *Journal of International Money and Finance*, 20, 249–272.

CHRISTENSEN, B. (2001): "Mismatch-Arbeitslosigkeit unter Geringqualifizierten," *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 34, 506–514.

CICCONE, A. AND R. E. HALL (1996): "Productivity and the Density of Economic Activity," *American Economic Review*, 86, 54–70.

CLIFF, A. D. AND J. K. ORD (1972): "Testing for spatial autocorrelation among regression residuals," *Geographical Analysis*, 4, 267–284.

——— (1981): *Spatial processes: Models and applications*, London: Pion.

COHEN, J. P. AND C. MORRISON PAUL (2007): "The impacts of transportation infrastructure on property values: a higher-order spatial econometrics approach," *Journal of Regional Science*, 47, 457–478.

COLES, M. G. AND E. SMITH (1996): "Cross-Section Estimation of the Matching Function: Evidence from England and Wales," *Economica*, 63, 589–597.

CRACOLICI, F. M., M. CUFFARO, AND P. NIJKAMP (2007): "Geographical Distribuition of Unemployment: An Analysis of Provincial Differences in Italy," *Growth and Change*, 38, 649–670.

DAVIES, S. AND M. HALLET (2001): "Policy responses to regional unemployment: Lessons from Germany, Spain and Italy," European Economy - Economic Papers 161, Directorate General Economic and Monetary Affairs, European Commission.

DEBARSY, N. AND C. ERTUR (2010): "Testing for spatial autocorrelation in a fixed effects panel data model," *Regional Science and Urban Economics*, 40, 453–470.

DMITRIJEVA, J. (2008): "Matching and labour market efficiency across space and through EU accesion: evidence from Latvia, Estonia and Slovenia," *EPEE Working Paper, Université d'Evry*.

ECKEY, H.-F., R. KOSFELD, AND M. TÜRCK (2007): "Regional Convergence in Germany: a Geographically Weighted Regression Approach," *Spatial Economic Analysis*, 2, 45–64.

ELHORST, J. P. (2003): “The Mystery of Regional Unemployment Differentials: theoretical and empirical explanations,” *Journal of Economic Surveys*, 17, 710–748.

——— (2010): “Dynamic Panels with Endogenous Interaction Effecrs when T is small,” *Regional Science and Urban Economics*, doi: 10.1016/j.regsciurbeco.2010.03.003.

——— (2012): “Dynamic spatial panel models: Models, methods and inferences,” *Journal of Geographical Systems*, 14, 5–28.

ELHORST, J. P., D. J. LACOMBE, AND G. PIRAS (2012): “On model specification and parameter space definitions in higher order spatial econometric models,” *Regional Science and Urban Economics*, 42, 211–220.

FAHR, R. AND U. SUNDE (2006): “Spatial mobility and competition for jobs: Some theory and evidence for Western Germany,” *Regional Science and Urban Economics*, 36, 803–825.

——— (2009): “Did the Hartz Reforms Speed-Up the Matching Process? A Macro-Evaluation Using Empirical Matching Functions,” *German Economic Review*, 10, 284–316.

FINGLETON, B. (1999): “Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration,” *Journal of Regional Science*, 39, 1–19.

FLORAX, R. J. G. M. AND S. REY (1995): “The Impacts of Misspecified Spatial Interaction in Linear Regression Models,” in *New directions in spatial econometrics*, ed. by L. Anselin and R. J. G. M. Florax, Berlin and and New York: Springer, 111–135.

FRANZESE, R. J. AND J. C. HAYS (2007): “Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data,” *Political Analysis*, 15, 140–164.

GRAU, A. (2009): “Pendler: Die Mehrheit nimmt weiter das Auto,” STATmagazin, Statistisches Bundesamt.

GRIFFITH, D. A. (1996): “Some Guidelines for Specifying the Geographic Weights Matrix Contained in Spatial Statistical Models,” in *Practical handbook of spatial statistics*, ed. by S. L. Arlinghaus, Boca Raton and Fla. [u.a.]: CRC Press, 65–82.

HALL, R. E. (1970): “Why is the Unemployment Rate so High at Full Employment?” *Brookings Papers on Economic Activity*, 1970, 369–410.

HUJER, R., P. J. M. RODRIGUES, AND K. WOLF (2009): “Estimating the macroeconomic effects of active labour market policies using spatial econometric methods,” *International Journal of Manpower*, 30, 648–671.

HYNNINEN, S.-M. (2005): "Matching across Space: Evidence from Finland," *LABOUR*, 19, 749–765.

IM, K. S., M. H. PESARAN, AND Y. SHIN (2003): "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 115, 53–74.

JAFFE, A. B. (1986): "Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value," *The American Economic Review*, 76, 984–1001.

KELEJIAN, H. H. AND I. R. PRUCHA (2010): "Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances," *Journal of Econometrics*, 157, 53–67.

KELEJIAN, H. H., I. R. PRUCHA, AND Y. YUZEFOVICH (2004): "Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small Sample Results," in *Spatial and spatiotemporal econometrics*, ed. by J. P. LeSage and R. K. Pace, Amsterdam: Elsevier JAI, 163–198.

KLOTZ, S. (2004): *Cross sectional dependence in spatial econometric models: With an application to German start-up activity data*, Münster: LIT.

KOPPEL, O. (2008): "Ingenieurarbeitsmarkt in Deutschland - gesamtwirtschaftliches Stellenangebot und regionale Fachkräftelücken," *IW-Trends – Vierteljahresschrift zur empirischen Wirtschaftsforschung aus dem Institut der deutschen Wirtschaft Köln*, 35, 1–16.

LACOMBE, D. J. (2004): "Does Econometric Methodology Matter? An Analysis of Public Policy Using Spatial Econometric Techniques," *Geographical Analysis*, 36, 105–118.

LEE, L.-F. AND X. LIU (2010): "Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances," *Econometric Theory*, 26, 187–230.

LEE, L.-F. AND J. YU (2010a): "A Spatial Dynamic Panel Data Model with Both Time and Individual Fixed Effects," *Econometric Theory*, 26, 564–597.

——— (2010b): "Estimation of spatial autoregressive panel data models with fixed effects," *Journal of Econometrics*, 154, 165–185.

——— (2010c): "Some recent developments in spatial panel data models," *Regional Science and Urban Economics*, 40, 255–271.

LEENDERS, R. T. A. J. (2002): "Modeling social influence through network autocorrelation: constructing the weight matrix," *Social Networks*, 24, 21–47.

LESAGE, J. P. AND R. K. PACE (2011): "Pitfalls in Higher Order Model Extensions of Basic Spatial Regression Methodology," *The Review of Regional Studies*, 41, 13–26.

LEVIN, A., C.-F. LIN, AND C.-S. J. CHU (2002): "Unit root tests in panel data: asymptotic and finite-sample properties," *Journal of Econometrics*, 108, 1–24.

LONGHI, S., P. NIJKAMP, AND J. POOT (2006): "Spatial Heterogeneity and the Wage Curve Revisited," *Journal of Regional Science*, 46, 707–731.

LOTTMANN, F. (2012a): "Explaining regional unemployment differences in Germany: a spatial panel data analysis," SFB 649 Discussion Paper 2012-026.

——— (2012b): "Spatial dependencies in German matching functions," *Regional Science and Urban Economics*, 42, 27–41.

MADDALA, G. S. AND S. WU (1999): "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics*, 61, 631–652.

MANNING, A. AND B. PETRONGOLO (2011): "How Local are Labor Markets?: Evidence from a Spatial Job Search Model," CEPR Discussion Paper No. 8686.

MARSTON, S. T. (1985): "Two Views of the Geographic Distribution of Unemployment," *Quarterly Journal of Economics*, 100, 57–79.

MARTIN, R. (1997): "Regional Unemployment Disparities and their Dynamics," *Regional Studies*, 31, 237–252.

MOLHO, I. (1995): "Spatial Autocorrelation in British Unemployment," *Journal of Regional Science*, 35, 641–658.

MÖLLER, J. AND A. ALDASHEV (2007): "Wage Inequality, Reservation Wages and Labor Market Participation: Testing the Implications of a Search-Theoretical Model with Regional Data," *International Regional Science Review*, 30, 120–151.

MORAN, P. A. P. (1950): "Notes on continuous phenomena," *Biometrika*, 37, 17–23.

NIEBUHR, A. (2003): "Spatial Interaction and Regional Unemployment in Europe," *European Journal of Spatial Development*.

NIEBUHR, A., N. GRANATO, A. HAAS, AND S. HAMANN (2012): "Doer Labour Mobility Reduce Disparities between Regional Labour Markets in Germany?" *Regional Studies*, 46, 841–858.

OCHSEN, C. (2009): "Regional Labor Markets and Aging in Germany," *Thünen-Series of Applied Economic Theory, No. 102, Universität Rostock*.

PÁEZ, A., D. M. SCOTT, AND E. VOLZ (2008): "Weight matrices for social influence analysis: An investigation of measurement errors and their effect on model identification and estimation quality," *Social Networks*, 30, 309–317.

PARENT, O. AND J. P. LESAGE (2008): "Using the Variance Structure of the Conditional Autoregressive Spatial Specification to Model Knowledge Spillovers," *Journal of Applied Econometrics*, 23, 235–256.

PARTRIDGE, M. D. AND D. S. RICKMAN (1997): "The Dispersion of US State Unemployment Rates: The Role of Market and Non-market Equilibrium Factors," *Regional Studies*, 31, 593–606.

PATACCHINI, E. AND Y. ZENOU (2007): "Spatial dependence in local unemployment rates," *Journal of Economic Geography*, 7, 169–191.

PERI, G. (2004): "Socio-Cultural Variables and Economic Success: Evidence from Italian Provinces 1951-1991," *Topics in Macroeconomics*, 4.

PESARAN, M. H. (2004): "General Diagnostic Tests for Cross Section Dependence in Panels," Cesifo working paper series no. 1229.

PETRONGOLO, B. AND C. A. PISSARIDES (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

ROSENTHAL, S. S. AND W. C. STRANGE (2001): "The Determinants of Agglomeration," *Journal of Urban Economics*, 50, 191–229.

STAKHIVYCH, S. AND T. H. A. BIJMOLT (2008): "Specification of spatial models: A simulation study on weights matrices," *Papers in Regional Science*, 88, 389–408.

STEVENS, M. (2007): "New Microfoundations for the Aggregate Matching Function," *International Economic Review*, 48, 847–868.

TAYLOR, J. AND S. BRADLEY (1997): "Unemployment in Europe: A Comparative Analysis of Regional Disparities in Germany, Italy and the UK," *KYKLOS*, 50, 221–245.

TIMM, U. (2008): "Wohnsituation in Deutschland 2006: Ergebnisse der Mikrozensus-Zusatzerhebung," *Wirtschaft und Statistik*, 312–321.

VAN OURS, J. C. (1991): "The Efficiency of the Dutch Labor Market in Matching Unemployment and Vacancies," *De Economist*, 139, 358–378.

YU, J., R. DE JONG, AND L.-F. LEE (2008): "Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large," *Journal of Econometrics*, 146, 118–134.

Appendix A

A.1 Panel unit root tests

A.1.1 Im, Pesaran and Shin (IPS) Test

Im et al. (2003) consider a sample of N cross sections observed over T time periods. They suppose that the stochastic process y_{it} is generated by a first-order autoregressive process:

$$y_{it} = (1 - \phi_i)\mu_i + \phi_i y_{i,t-1} + \epsilon_{it}, \qquad i = 1, \ldots, N, t = 1, \ldots, T \tag{A.1.1}$$

where initial values y_{i0} are given. To test the null hypothesis of unit roots, i.e. $\phi_i = 1$ for all i, Im et al. (2003) express equation (A.1.1) further:

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \epsilon_{it} \tag{A.1.2}$$

where $\alpha_i = (1 - \phi_i)\mu_i$, $\beta_1 = -(1 - \phi_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$. In this formulation, the null hypothesis that each series in the panel contains a unit root and the alternative allowing for some (but not all) of the individual series to have a unit root, correspond to

$$H_0 : \beta_i = 0 \, \forall i \tag{A.1.3}$$

and

$$H_1 : \begin{cases} \beta_i < 0 & i = 1, \ldots, N_1 \\ \beta_i = 0 & i = N_1 + 1, \ldots, N. \end{cases} \tag{A.1.4}$$

This formulation of the alternative is more general than the homogeneous alternative, i.e. $\beta_i = \beta < 0$. Im et al. (2003) assume that under the alternative hypothesis the fraction of the individual processes that are stationary is nonzero, namely if $\lim\limits_{N \to \infty} (N_1/N) = \delta, 0 < \delta \leq 1$. This condition is necessary for the consistency of the test. Im et al. (2003) propose both unit root tests for heterogeneous panels with fixed T and serially uncorrelated errors and unit root tests for heterogeneous panels with serially correlated errors. For the sake of brevity, we only consider the test for serially uncorrelated errors. The IPS t-bar statistic is defined as the average of the individual ADF statistics, i.e.

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{iT} \tag{A.1.5}$$

where t_{iT} is the individual t-statistic for testing $H_0 : \beta_i = 0$ for all i in equation (A.1.4). Im et al. (2003) show that for heterogeneous panels with serially uncorrelated errors the standardized t-bar statistic is distributed as standard normal as $N \to \infty$ for a fixed T, as long as $T > 5$ in the case of DF regressions with intercepts and $T > 6$ in the case of DF regressions with intercepts and linear time trends. Finally, in Monte Carlo experiments, Im et al. (2003) show that if a large enough lag order is selected for the underlying ADF regressions, then the small sample performance of the t-bar test is reasonably satisfactory and generally better than the Levin et al. (2002) (LLC) test.

Fisher-type Tests

Let G_{iT_i} be a unit root test statistic for the ith group in a panel and assume that as the time series observations for the ith group $T_i \to \infty$, $G_{iT_i} \Rightarrow G_i$ where G_i is a nondegenerate random variable. Let p_i be the asymptotic p-value of a unit root test for cross-section i, i.e. $p_i = F(G_{iT_i})$, where $F(\bullet)$ is the distribution function of the random variable G_i (Baltagi (2005)). Maddala and Wu (1999) propose a Fisher-type test

$$P = -2 \sum_{i=1}^{N} ln p_i \tag{A.1.6}$$

which combines the p-values from unit root tests for each cross-section i to test for unit roots in the panel data set. The statistic P has a χ^2 distribution with two degrees of freedom as $T_i \to \infty$ for finite N. Maddala and Wu (1999) argue that the advantage of this test is, firstly, that no balanced panel is required as it is the case for the IPS test. Secondly, it is possible to use different lag lengths in the individual ADF regressions and, thirdly, it can be carried out for any panel unit root test. However, the p-values have to be derived by Monte Carlo simulation which is a disadvantage of this test. Moreover, Maddala and Wu (1999) find that the Fisher-type test with bootstrap-based critical values performs the best and is the preferred choice for testing the null hypothesis of nonstationarity as well as in testing for cointegration in panels.

Appendix B

B.1 District reforms

There are two district reforms during the period from 1999 until 2009. In the federal state of Saxony-Anhalt, the reform became effective as from July 1st, 2007. The reform reduced the number of districts from 21 to 11. The other reform affects the federal state of Saxony where it became effective as from August 1st, 2008. The reform resulted in a reduced number of districts from 29 to 13. As we want to use current map data of Germany, we aggregate the data according to the reforms. Hence, we use the regional structure after the reforms for the entire period. Regional unemployment rates of the 'new' district are weighted averages of the corresponding regional unemployment rates using the associated labor force as weights.

B.2 Summary statistics of explanatory variables

	Min	1st Qu.	Median	Mean	3rd Qu.	Max	Std. dev.	unit
EG	-4.084	-0.859	-0.002	-0.234	0.537	2.724	1.178	%
$\%IND$	3.455	19.580	26.350	27.650	35.990	67.240	11.254	%
$\%ENERW$	0	0.478	0.739	0.904	1.074	7.974	0.732	%
$\%CON$	2.130	5.681	7.559	7.938	10.1	17.4	3.035	%
$\%HOT$	0.849	2.002	2.407	2.97	3.148	22.08	1.989	%
$\%FIN$	0.445	2.096	2.612	2.941	3.232	16.95	1.73	%
$\%PUB$	2.334	4.882	6.012	6.742	8.136	19.32	2.679	%
$YOUNG$	13.71	16.53	17.5	17.54	18.43	23.36	1.45	%
OLD	21.46	26.73	27.93	27.91	29.17	33.71	1.88	%
$H0$	7.351	14.44	17.91	17.1	20.32	30.57	4.63	%
$H1$	50.1	61.3	64.38	64.35	67.25	77.93	4.983	%
REG	0.335	0.923	1.404	1.992	2.258	39.27	2622	thousand
$DEBTR$	0.025	3.002	4.363	4.71	6.127	13.49	2.24	ratio

Table B.1 Summary statistics of the explanatory variables of our model using averages over the period from 1999 until 2007

B.3 Specification test by Debarsy and Ertur (2010)

Consider the SARAR(1,1) model[1]

$$Y_t = \lambda W Y_t + X_t\beta + \mu + U_t; \quad U_t = \rho M U_t + V_t; \quad t = 1, \ldots, T, \tag{B.3.1}$$

where $Y_t = (y_{1,t}, y_{2,t}, \ldots, y_{n,t})'$ is the $(n \times 1)$ vector of the dependent variable for all individuals in period t, X_t is the $(n \times k)$ matrix of exogenous regressors and β is the associated vector of unknown regression coefficients. $V_t = (v_{1,t}, \ldots, v_{n,t})'$ is the innovation term where $v_{i,t}$ is i.i.d. across i and t with zero mean and constant variance σ^2. μ is the $(n \times 1)$ vector of individual effects. W and M are $(n \times n)$ spatial weights matrices, λ and ρ are the unknown spatial parameters. Applying the transformation approach by Lee and Yu yields the transformed model

$$Y_t^* = \lambda W Y_t^* + X_t^*\beta + U_t^*; \quad U_t^* = \rho M U_t^* + V_t^*; \quad t = 1, \ldots, T-1. \tag{B.3.2}$$

Denoting $\theta' = [\beta', \rho, \lambda, \sigma^2]$ and $\eta' = [\beta', \rho, \lambda]$, the log-likelihood function is given by

$$l(\theta) = -\frac{n(T-1)}{2} ln(2\pi) - \frac{n(T-1)}{2} ln(\sigma^2) + (T-1) ln|S(\lambda)| + (T-1) ln|R(\rho)| - \frac{1}{2\sigma^2}\sum_{t=1}^{T-1} V^{*\prime}_t(\eta) V_t^*(\eta), \tag{B.3.3}$$

where $S(\lambda) = I_n - \lambda W$, $R(\rho) = I_n - \rho M$ and $V_t^* = R(\rho)[S(\lambda)Y_t^* - X_t^*\beta]$. Debarsy and Ertur (2010) consider five different hypotheses (H_0^a until H_0^e) for their specification test.

Joint test statistic for $H_0^a : \rho = \lambda = 0$

Under the null hypothesis, the specification to be estimated is the standard fixed effects panel model. Debarsy and Ertur (2010) show that the transformed model can also be estimated by OLS. The joint LM statistic is then given by

$$LM_a = \tilde{Q}^{-1}[T_{22}\tilde{R}_y^2 - 2T_{12}\tilde{R}_v\tilde{R}_y + (\tilde{D} + T_{11})\tilde{R}_v^2]; \tag{B.3.4}$$

where

$$\tilde{R}_v = \frac{\sum_{t=1}^{T-1} \tilde{V}_t^{*\prime} M \tilde{V}_t^*}{\tilde{\sigma}^2};$$

$$\tilde{R}_y = \frac{\sum_{t=1}^{T-1} \tilde{V}_t^{*\prime} W Y_t^*}{\tilde{\sigma}^2};$$

$$\tilde{D} = \tilde{\sigma}^{-2}\sum_{t=1}^{T-1} (WX_t^*\tilde{\beta})' M_{X^*} (WX_t^*\tilde{\beta});$$

[1] All explanations are taken from the original paper by Debarsy and Ertur (2010).

$$\tilde{Q} = (\tilde{D} + T_{11})T_{22} - T_{12}^2.$$

Also, $T_{11} = (T-1)tr[(W+W')W]$, $T_{22} = (T-1)tr[(M+M')M]$, $T_{12} = (T-1)tr((M'+M)W)$ and $M_{X^*} = I_n - X_t^*({X^*}'_t X_t^*)^{-1}{X^*}'_t$. Finally, $\tilde{V}_t^* = Y_t^* - X_t^*\tilde{\beta}$ is the residual of the constrained model and $\tilde{\sigma}^2$ is the associated OLS residual variance. LM_a is expected to be asymptotically distributed as χ_2^2 under the joint null hypothesis H_0^a.

Marginal test statistic for $H_0^b : \lambda = 0$ (assuming $\rho = 0$)

Under this hypothesis, the constrained model is

$$Y_t^* = X_t^*\beta + V_t^*; \quad t = 1, \ldots, T-1, \tag{B.3.5}$$

where V_t^* is distributed according to a multivariate normal distribution with zero mean and the variance-covariance matrix $\sigma^2 I_{n(T-1)}$. The unconstrained model is

$$Y_t^* = \lambda W Y_t^* + X_t^*\beta + V_t^*; \quad t = 1, \ldots, T-1, \tag{B.3.6}$$

and should be estimated using maximum likelihood. The LM statistic for this hypothesis is

$$LM_b = \frac{\sum_{t=1}^{T-1}(\tilde{V}_t^{*\prime} W Y_t^* / \tilde{\sigma}^2)^2}{\tilde{D} + T_{11}}, \tag{B.3.7}$$

where the variables are defined as before. $\tilde{V}_t^*$ are the OLS residuals of equation (B.3.5). This LM statistic is asymptotically distributed as χ_1^2 under the null hypothesis.

Marginal test statistic for $H_0^c : \rho = 0$ (assuming $\lambda = 0$)

For this test, the restricted model is again equation (B.3.5). The specification under the alternative is

$$Y_t^* = X_t^*\beta + U_t^*; \quad U_t^* = \rho M U_t^* + V_t^*; \quad t = 1, \ldots, T-1. \tag{B.3.8}$$

The LM statistic for this hypothesis is

$$LM_c = \frac{\sum_{t=1}^{T-1}(\tilde{V}_t^{*\prime} M \tilde{V}_t^* / \tilde{\sigma}^2)^2}{T_{22}}. \tag{B.3.9}$$

Again, $\tilde{V}_t^*$ are the residuals of equation (B.3.5) and $\tilde{\sigma}^2$ is the estimate of the corresponding residual variance. Under the null hypothesis, LM_c is asymptotically distributed as χ_1^2.

Conditional test statistic for $H_0^d : \rho = 0$ given $\lambda \neq 0$

The appropriate specification under the null is model (B.3.6). When the null is rejected, the correct specification is the general model (B.3.2). The disturbances of the restricted model are given by

$$V_t^* = S(\lambda)Y_t^* - X_t^*\beta; \quad t = 1, \ldots, T-1. \tag{B.3.10}$$

These disturbances can be estimated by ML in model (B.3.6). The *LM* statistic for the conditional test for spatially autocorrelated errors in the presence of an endogenous spatial lag is given by

$$LM_d = \frac{(\sum_{t=1}^{T-1} \tilde{V}_t^{*\prime} M \tilde{V}_t^* / \tilde{\sigma}^2)^2}{T_{22} - (\tilde{T}_{\lambda\rho})^2 * var(\tilde{\rho})}, \tag{B.3.11}$$

where $var(\tilde{\rho})$ is the variance of the autoregressive coefficient estimated under the constrained model and $\tilde{T}_{\lambda\rho} = (T-1)tr[M'WS(\tilde{\rho})^{-1} + MWS(\tilde{\rho})^{-1}]$. LM_d is asymptotically distributed as χ_1^2 under the null hypothesis.

Conditional test statistic for $H_0^e : \lambda = 0$ given $\rho \neq 0$

For this test, the unconstrained model is the general specification (equation B.3.2) whereas the constrained model is equation (B.3.8). Its error term is given by

$$V_t^* = R(\rho)[Y_t^* - X_t^*\beta]; \quad t = 1, \ldots, T-1. \tag{B.3.12}$$

The conditional *LM* statistic for the conditional test for an endogenous spatial lag in the presence of spatially autocorrelated errors is

$$LM_e = \frac{\sum_{t=1}^{T-1} (\tilde{V}_t^{*\prime} R(\tilde{\rho}) W Y_t^* / \tilde{\sigma}^2)^2}{\tilde{I}_{11} - \tilde{I}_{12}\tilde{I}_{22}\tilde{I}_{21}} \tag{B.3.13}$$

with $\tilde{I}_{22}$ being the variance-covariance matrix of the non-constrained parameters, namely $\tilde{\rho}$, $\tilde{\beta}$ and $\tilde{\sigma}^2$. The other terms are defined by

$$\tilde{I}_{11} = (T-1)tr(W)^2 + \frac{1}{\tilde{\sigma}^2} \sum_{t=1}^{T-1} [(R(\tilde{\rho})WX_t^*\tilde{\beta})'(R(\tilde{\rho})WX_t^*\tilde{\beta})] \\ + (T-1)tr[(R(\tilde{\rho})WR(\tilde{\rho})^{-1})'(R(\tilde{\rho})WR(\tilde{\rho})^{-})]; \tag{B.3.14}$$

$$\tilde{I}_{12} = \begin{bmatrix} \frac{1}{\tilde{\sigma}^2} \sum_{t=1}^{T-1} X_t^{*\prime} R(\tilde{\rho})' R(\tilde{\rho}) W X_t^* \tilde{\beta} \\ (T-1)tr[(MR(\tilde{\rho})^{-1})'R(\tilde{\rho})WR(\tilde{\rho})^{-1} + MWR(\tilde{\rho})^{-1}] \\ 0 \end{bmatrix}. \tag{B.3.15}$$

All parameters involved in this test come from the contrained model. The LM_e statistic is asymptotically distributed as χ_1^2.

B.4 Test for serial correlation, spatial autocorrelation and random effects by Baltagi et al. (2007b)

Consider the panel data model[2]

$$y_{ti} = x'_{ti}\beta + u_{ti}; \quad i = 1, \ldots, n; \quad t = 1, \ldots, T, \tag{B.4.16}$$

where y_{ti} is the observation of the ith region for the tth time period, x_{ti} denotes the $(k \times 1)$ vector of observations on the nonstochastic regressors and u_{ti} is the regression disturbance. In vector form, the disturbance vector of equation (B.4.16) is assumed to have random region effects, spatially autocorrelated residual disturbances and a first-order autoregressive remainder disturbance term and it is written by

$$u_t = \mu + \epsilon_t \tag{B.4.17}$$

with

$$\epsilon_t = \lambda W \epsilon_t + v_t \tag{B.4.18}$$

and

$$v_t = \rho v_{t-1} + e_t, \tag{B.4.19}$$

where $u'_t = (u_{t1}, \ldots, u_{tN})$ and ϵ_t, v_t and e_t are similarly defined. $\mu = (\mu_1, \mu_2, \ldots, \mu_N)'$ denote the vector of random region effects which are assumed to be $IIN(0, \sigma^2_\mu)$. λ is the scalar spatial autoregressive coefficient while ρ is the time-wise serial correlation coefficient. W is the $(n \times n)$ spatial weights matrix. $e_{ti} \sim IIN(0, \sigma^2_e)$ and $v_{i,0} \sim N((0, \sigma^2_e)/(1-\rho^2))$.

Baltagi et al. (2007b) consider joint, marginal and conditional hypotheses for spatial error dependence, random region effects as well as for serial error dependence. Table B.2 presents an overview of the hypotheses, test statistics and asymptotic distributions. To save space, we define all terms only once in the table.

[2] All explanations are taken from the original paper by Baltagi et al. (2007b).

	Joint Test	One-dimensional marginal tests		
Hypothesis	$\rho = \lambda = \sigma_\mu^2 = 0$	$\lambda = 0$ (assuming $\rho = \sigma_\mu^2 = 0$)	$\rho = 0$ (assuming $\lambda = \sigma_\mu^2 = 0$)	$\sigma_\mu^2 = 0$ (assuming $\rho = \lambda = 0$)
LM statistic	$\mathrm{LM}_J = \frac{nT^2}{2(T-1)(T-2)}\left[A^2 - 4AF + 2TF^2\right] + \frac{n^2T}{b}H^2$ where $A = \frac{\tilde{u}'(J_T \otimes I_n)\tilde{u}}{\tilde{u}'\tilde{u}} - 1$ $F = \frac{1}{2}\left(\frac{\tilde{u}'(G \otimes I_n)\tilde{u}}{\tilde{u}'\tilde{u}}\right)$ $H = \frac{1}{2}\left(\frac{\tilde{u}'(I_T \otimes (W'+W))\tilde{u}}{\tilde{u}'\tilde{u}}\right)$ $b = \mathrm{tr}(W + W')^2/2 = \mathrm{tr}(W^2 + W'W)$ $\tilde{u}$ = OLS residuals G = bidiagonal matrix with bidiagonal elements equal to one J_T is a matrix of ones of dimension T	$\mathrm{LM}_\lambda = \frac{n^2T}{b}H^2$	$\mathrm{LM}_\rho = \frac{nT^2}{T-1}F^2$	$\mathrm{LM}_\mu = \frac{nT}{2(T-1)}A^2$
asymptotic distribution under H_0	χ_3^2	χ_1^2	χ_1^2	χ_1^2

	Two-dimensional marginal tests			
Hypothesis	$\sigma_\mu^2 = \rho = 0$ (assuming $\lambda = 0$)	$\lambda = \rho = 0$ (assuming $\sigma_\mu^2 = 0$)	$\lambda = \sigma_\mu^2 = 0$ (assuming $\rho = 0$)	
LM statistic	$\mathrm{LM}_{\mu\rho} = \frac{nT^2}{2(T-1)(T-2)}\left[A^2 - 4AF + 2TF^2\right]$	$\mathrm{LM}_{\lambda\rho} = \mathrm{LM}_\lambda + \mathrm{LM}_\rho$	$\mathrm{LM}_{\lambda\mu} = \mathrm{LM}_\lambda + \mathrm{LM}_\mu$	
asymptotic distribution under H_0	χ_2^2	χ_2^2	χ_2^2	

	One-dimensional conditional tests
Hypothesis	$\lambda = 0 \quad$ (allowing $\rho \neq 0$ and $\sigma_\mu^2 > 0$)
LM statistic	$\mathrm{LM}_{\lambda/\rho\mu} = \frac{\widehat{D}(\lambda)^2}{b(T-2cg+c^2g^2)}$ where $g = \mathrm{tr}(V^{-1}J_T) = \frac{1}{\sigma_e^2}(1-\rho)\{2+(T-2)(1-\rho)\}$ $c = \frac{\sigma_e^2\sigma_\mu^2}{d^2(1-\rho)^2\sigma_\mu^2+\sigma_e^2}$ $\widehat{D}(\lambda) = \frac{\partial L}{\partial \lambda}\big\vert_{H_0} = \frac{1}{2}\hat{u}'\Big(V^{-1} - 2cV^{-1}J_TV^{-1} + c^2[V^{-1}J_T]^2V^{-1}\Big) \otimes (W'+W)\hat{u}$ $\hat{u}$ are the restricted maximum likelihood residuals under H_0 and $V^{-1} = \frac{1}{\sigma_e^2}V_\rho^{-1} \quad$ with $\quad V_\rho = \frac{1}{1-\rho^2}V_1 \quad$ where $V_1 = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{pmatrix}$
asymptotic distribution under H_0	χ_1^2

	One-dimensional conditional tests continued
Hypothesis	$\rho = 0 \quad$ (allowing $\lambda \neq 0$ and $\sigma_\mu^2 > 0$)
LM statistic	$\mathrm{LM}_{\rho/\lambda\mu} = \widehat{D}^2(\rho)\, J_{33}^{-1}$ where $\widehat{D}(\rho) = \frac{\partial L}{\partial \rho}\big\vert_{H_0}$ $= -\frac{(T-1)}{T}\big(\widehat{\sigma}_e^2\, \mathrm{tr}(Z_0(B'B)^{-1}) - n\big) + \frac{1}{2}\widehat{\sigma}_e^2 \hat{u}'\Big(\frac{1}{\widehat{\sigma}_e^4}(E_T\, G\, E_T) \otimes (B'B) + \frac{1}{\widehat{\sigma}_e^2}(\bar{J}_T\, G\, E_T) \otimes Z_0 + \frac{1}{\widehat{\sigma}_e^2}(E_T\, G\, \bar{J}_T) \otimes Z_0 + (\bar{J}_T\, G\, \bar{J}_T) \otimes$ $Z_0(B'B)^{-1}Z_0\Big)\hat{u}$ with $Z_0 = \big(T\sigma_\mu^2 I_n + \sigma_e^2(B'B)^{-1}\big)^{-1}, \quad B = I_n - \lambda W, \quad \bar{J}_T = J_T/T \quad , J_T = \mathbb{1}_T\mathbb{1}_T' \quad$ and $\quad E_T = I_T - \bar{J}_T$ J_{33}^{-1} is the (3,3) element of the inverse of the information matrix $\widehat{J}_\theta$ evaluated under H_0. $\widehat{J}_\theta = \begin{pmatrix} \frac{1}{2}\big(\frac{n(T-1)}{\widehat{\sigma}_e^4} + d_1\big) & \frac{T}{2}d_2 & \frac{(T-1)}{T}\big(\widehat{\sigma}_e^2 d_1 - \frac{n}{\widehat{\sigma}_e^2}\big) & \frac{1}{2}\big[\frac{(T-1)}{\widehat{\sigma}_e^2}d_3 + \widehat{\sigma}_e^2 d_4\big] \\ \frac{T}{2}d_2 & \frac{T^2}{2}\mathrm{tr}[Z_0]^2 & (T-1)\widehat{\sigma}_e^2 d_2 & \frac{T\widehat{\sigma}_e^2}{2}d_5 \\ \frac{T-1}{T}\big(\widehat{\sigma}_e^2 d_1 - \frac{n}{\widehat{\sigma}_e^2}\big) & (T-1)\widehat{\sigma}_e^2 d_2 & \widehat{J}_{\rho\rho} & \frac{T-1}{T}\big(\widehat{\sigma}_e^4 d_4 - d_3\big) \\ \frac{1}{2}\big(\frac{T-1}{\widehat{\sigma}_e^2}d_3 + \widehat{\sigma}_e^2 d_4\big) & \frac{T\widehat{\sigma}_e^2}{2}d_5 & \frac{T-1}{T}\big(\widehat{\sigma}_e^4 d_4 - d_3\big) & \frac{1}{2}\big[(T-1)d_6 + \widehat{\sigma}_e^4 d_7\big] \end{pmatrix}$ with $d_1 = \mathrm{tr}\,\big(Z_0(B'B)^{-1}\big)^2$ $d_2 = \mathrm{tr}\,\big(Z_0(B'B)^{-1}Z_0\big)$ $d_3 = \mathrm{tr}\,\big((W'B + B'W)(B'B)^{-1}\big)$ $d_4 = \mathrm{tr}\,\big(Z_0(B'B)^{-1}(W'B + B'W)(B'B)^{-1}Z_0(B'B)^{-1}\big)$ $d_5 = \mathrm{tr}\,\big(Z_0(B'B)^{-1}(W'B + B'W)(B'B)^{-1}Z_0\big)$ $d_6 = \mathrm{tr}\,\big((W'B + B'W)(B'B)^{-1}\big)^2$ $d_7 = \mathrm{tr}\,\big(Z_0(B'B)^{-1}(W'B + B'W)(B'B)^{-1}\big)^2$ and $\widehat{J}_{\rho\rho} = \frac{n}{T^2}(T^3 - 3T^2 + 2T + 2) + \frac{2(T-1)^2\widehat{\sigma}_e^4}{T^2}d_1$
asymptotic distribution under H_0	χ_1^2

	One-dimensional conditional tests continued
Hypothesis	$\sigma_\mu^2 = 0$ (allowing $\rho \neq 0$ and $\lambda \neq 0$)
LM statistic	$$\text{LM}_{\mu/\rho\lambda} = \widehat{D}(\sigma_\mu^2) J_{22}^{-1}$$ where $\widehat{D}(\sigma_\mu^2) = \frac{\partial L}{\partial \sigma_\mu^2}\big\vert_{H_0} = -\frac{g}{2}\text{tr}(B'B) + \frac{1}{2\sigma_e^4}\hat{u}'\left[V_\rho^{-1} J_T V_\rho^{-1} \otimes (B'B)^2\right]\hat{u}$ with $\hat{u}$ = restricted maximum likelihood residuals under H_0 J_{22}^{-1} is the (2,2) element of the inverse of the information matrix $\widehat{J}_\theta$ evaluated under H_0. $$\widehat{J}_\theta = \begin{pmatrix} \frac{nT}{2\sigma_e^4} & \frac{g\,\text{tr}(B'B)}{2\sigma_e^2} & \frac{n\rho}{\sigma_e^2(1-\rho^2)} & \frac{Td_3}{2\sigma_e^2} \\ \frac{g\,\text{tr}(B'B)}{2\sigma_e^2} & \frac{g^2\text{tr}(B'B)^2}{2} & \frac{\text{tr}(B'B)}{\sigma_e^2(1+\rho)}[(2-T)\rho^2+(T-1)+\rho] & \frac{g}{2}\text{tr}[W'B+B'W] \\ \frac{n\rho}{\sigma_e^2(1-\rho^2)} & \frac{\text{tr}(B'B)}{\sigma_e^2(1+\rho)}[(2-T)\rho^2+(T-1)+\rho] & \frac{n}{(1-\rho^2)^2}[3\rho^2-\rho^2T+T-1] & \frac{\rho d_3}{1-\rho^2} \\ \frac{Td_3}{2\sigma_e^2} & \frac{g}{2}\text{tr}[W'B+B'W] & \frac{\rho d_3}{1-\rho^2} & \frac{Td_6}{2} \end{pmatrix}$$
asymptotic distribution under H_0	χ_1^2

	Two-dimensional conditional tests
Hypothesis	$\lambda = \rho = 0$ (allowing $\sigma_\mu^2 > 0$)
LM statistic	$\mathrm{LM}_{\lambda\rho/\mu} = \widehat{D}_\theta' \widehat{J}_\theta^{-1} \widehat{D}_\theta = \frac{\widehat{D}(\rho)^2 n^2 T^2 (T-1)}{4\widehat{\sigma}_1^4 \widehat{\sigma}_e^4 \det[J(\theta_1)]} + \frac{\widehat{D}(\lambda)^2}{\left[(T-1) + \frac{\widehat{\sigma}_e^4}{\widehat{\sigma}_1^4}\right] b}$ with $\widehat{D}(\rho) = \frac{\partial L}{\partial \rho}\big\vert_{H_0} = \frac{n(T-1)}{T}\left(\frac{\widehat{\sigma}_1^2 - \widehat{\sigma}_e^2}{\widehat{\sigma}_1^2}\right) + \frac{1}{2}\widehat{\sigma}_e^2 \hat{u}' \left[(\bar{J}_T/\widehat{\sigma}_1^2 + E_T/\widehat{\sigma}_e^2)\, G\, (\bar{J}_T/\widehat{\sigma}_1^2 + E_T/\widehat{\sigma}_e^2) \otimes I_n \right] \hat{u}$ $\widehat{D}(\lambda) = \frac{\partial L}{\partial \lambda}\big\vert_{H_0} = \frac{1}{2}\hat{u}' \left[\frac{\widehat{\sigma}_e^2}{\widehat{\sigma}_1^4} \left(\bar{J}_T \otimes (W' + W)\right) + \frac{1}{\widehat{\sigma}_e^2}\left(E_T \otimes (W' + W)\right) \right] \hat{u}$ where $\widehat{\sigma}_1^2 = \hat{u}'(J_T \otimes I_n)\hat{u}/nT$ and $\widehat{\sigma}_e^2 = \hat{u}'(E_T \otimes I_n)\hat{u}/n(T-1)$ are the solutions of $\frac{\partial L}{\partial \sigma_\mu^2}\big\vert_{H_0} = 0$ and $\frac{\partial L}{\partial \sigma_e^2}\big\vert_{H_0} = 0$, respectively. $\hat{u}$ are the restricted maximum likelihood residuals under H_0 $J(\theta_1)$ is the block-diagonal information matrix with respect to the parameters $(\sigma_e^2, \sigma_\mu^2, \rho)$. The information matrix is given by $\widehat{J}_\theta = \begin{pmatrix} \frac{n}{2}\left(\frac{1}{\widehat{\sigma}_1^4} + \frac{T-1}{\widehat{\sigma}_e^4}\right) & \frac{nT}{2\widehat{\sigma}_1^4} & \frac{n(T-1)}{T}\widehat{\sigma}_e^2\left(\frac{1}{\widehat{\sigma}_1^4} - \frac{1}{\widehat{\sigma}_e^4}\right) & 0 \\ \frac{nT}{2\widehat{\sigma}_1^4} & \frac{nT^2}{2\widehat{\sigma}_1^4} & \frac{n(T-1)\widehat{\sigma}_e^2}{\widehat{\sigma}_1^4} & 0 \\ \frac{n(T-1)}{T}\widehat{\sigma}_e^2\left(\frac{1}{\widehat{\sigma}_1^4} - \frac{1}{\widehat{\sigma}_e^4}\right) & \frac{n(T-1)\widehat{\sigma}_e^2}{\widehat{\sigma}_1^4} & \widehat{J}_{\rho\rho} & 0 \\ 0 & 0 & 0 & (T-1)b + \frac{\widehat{\sigma}_e^4}{\widehat{\sigma}_1^4} b \end{pmatrix}$ where $\widehat{J}_{\rho\rho} = n\left[2a^2(T-1)^2 + 2a(2T-3) + T - 1\right]$ with $a = \frac{\widehat{\sigma}_e^2 - \widehat{\sigma}_1^2}{T\widehat{\sigma}_1^2}$
asymptotic distribution under H_0	χ_2^2

	Two-dimensional conditional tests continued
Hypothesis	$\lambda = \sigma_\mu^2 = 0$ (allowing $\rho \neq 0$)
LM statistic	$$\mathrm{LM}_{\lambda\mu/\rho} = \widehat{D}_\theta' \widehat{J}_\theta^{-1} \widehat{D}_\theta = \frac{\widehat{D}^2(\sigma_\mu^2)}{\det[J(\theta_1)]} \frac{n^2}{\sigma_e^4(1-\rho^2)} \left\{ \frac{T}{2}(3\rho^2 - \rho^2 T + T - 1) - \rho^2 \right\} + \frac{\widehat{D}^2(\lambda)}{Tb}$$ where $\widehat{D}(\sigma_\mu^2) = \frac{\partial L}{\partial \sigma_\mu^2}\big\rvert_{H_0} = -\frac{ng}{2} + \frac{1}{2\sigma_e^4} \hat{u}' \left(V_\rho^{-1} J_T V_\rho^{-1} \otimes I_n \right) \hat{u}$ $\widehat{D}(\lambda) = \frac{\partial L}{\partial \lambda}\big\rvert_{H_0} = \frac{1}{2\sigma_e^2} \hat{u}' \left(V_\rho^{-1} \otimes (W' + W) \right) \hat{u}$ The information matrix is given by $$\widehat{J}_\theta = \begin{pmatrix} \frac{nT}{2\sigma_e^4} & \frac{ng}{2\sigma_e^2} & \frac{n\rho}{\sigma_e^2(1-\rho^2)} & 0 \\ \frac{ng}{2\sigma_e^2} & \frac{ng^2}{2} & \frac{n}{\sigma_e^2(1+\rho)} \left[(2-T)\rho^2 + \rho + (T-1) \right] & 0 \\ \frac{n\rho}{\sigma_e^2(1-\rho^2)} & \frac{n}{\sigma_e^2(1+\rho)} \left[(2-T)\rho^2 + \rho + (T-1) \right] & \frac{n}{(1-\rho^2)^2}(3\rho^2 - \rho^2 T + T - 1) & 0 \\ 0 & 0 & 0 & Tb \end{pmatrix}$$ $\hat{u}$ are the restricted maximum likelihood residuals under H_0 $J(\theta_1)$ is the block-diagonal information matrix with respect to the parameters $(\sigma_e^2, \sigma_\mu^2, \rho)$.
asymptotic distribution under H_0	χ_2^2

	Two-dimensional conditional tests continued
Hypothesis	$\sigma_\mu^2 = \rho = 0$ (allowing $\lambda \neq 0$)
LM statistic	$$\mathrm{LM}_{\mu\rho/\lambda} = \widehat{D}_\theta' \widehat{J}_\theta^{-1} \widehat{D}_\theta$$ where $\widehat{D}_\theta' = \left(0, \widehat{D}(\sigma_\mu^2), \widehat{D}(\rho), 0\right)$ with $\widehat{D}(\sigma_\mu^2) = \frac{\partial L}{\partial \sigma_\mu^2}\big\rvert_{H_0} = -\frac{T}{2\sigma_e^2}\mathrm{tr}(B'B) + \frac{1}{2\sigma_e^4}\hat{u}'\left(J_T \otimes (B'B)^2\right)\hat{u}$ $\widehat{D}(\rho) = \frac{\partial L}{\partial \rho}\big\rvert_{H_0} = \frac{1}{2\sigma_e^2}\hat{u}'\left[G \otimes (B'B)\right]\hat{u}$ The information matrix is given by $\widehat{J}_\theta = \begin{pmatrix} \frac{nT}{2\sigma_e^4} & \frac{T}{2\sigma_e^4}\mathrm{tr}(B'B) & 0 & \frac{T}{2\sigma_e^4}d_3 \\ \frac{T}{2\sigma_e^4}\mathrm{tr}(B'B) & \frac{T^2}{2\sigma_e^4}\mathrm{tr}\left[(B'B)^2\right] & \frac{T-1}{\sigma_e^2}\mathrm{tr}(B'B) & \frac{T}{2\sigma_e^4}\mathrm{tr}[W'B + B'W] \\ 0 & \frac{T-1}{\sigma_e^2}\mathrm{tr}(B'B) & n(T-1) & 0 \\ \frac{T}{2\sigma_e^4}d_3 & \frac{T}{2\sigma_e^4}\mathrm{tr}[W'B + B'W] & 0 & \frac{T}{2\sigma_e^4}d_6 \end{pmatrix}$ $\hat{u}$ are the restricted maximum likelihood residuals under H_0
asymptotic distribution under H_0	χ_2^2

Table B.2 Hypotheses, LM statistics and asymptotic distributions of Baltagi et al. (2007b)

B.5 Data description

Variable	Description	Data Source
	dependent variable	
unemployment rate (u_{it})	Regional unemployment rate at district level. Unemployment is defined according to regulations in Social Security Code, i.e. a person is officially registered as unemployed when certain requirements are fulfilled as this status is connected to the right to receive public benefits.	Federal Employment Office (Bundesagentur fr Arbeit), data can be downloaded from regional data base of Federal Statistical Office (`www.regionalstatistik.de`).
	explanatory variables	
market equilibrium effects		
Industry mix ($\%AGR_{it}$, $\%IND_{it}$, ...)	Number of employed individuals which have a job that is subject to social insurance contribution according to the sector in which they are working in relation to the total number of employed individuals with a job that is subject to social contributions. The classification of sectors in the version of 2003 is used which bases upon the European classification (NACE Rev. 1.1)). As there is a change in the sector classification in 2008, we only use data until 2007.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`). Some values for the federal country of Saxony-Anhalt are missing in this database. We requested them directly from the regional statistical office of Saxony-Anhalt.
demographic variables		
Age structure of the population (OLD_{it}, $YOUNG_{it}$)	Share of labor force aged 15 until 25 and older than 50 years, respectively. The labor force consists of employed and unemployed individuals.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`). Some values for the federal country of Saxony-Anhalt and Saxony are missing in this database. We requested them directly from the regional statistical office of Saxony-Anhalt and Saxony.

Variable	Description	Data Source
Foreigners ($FOREIGN_{it}$)	Extent of foreign labor force in relation to the whole labor force in district *i*.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`). Some values for the federal country of Saxony-Anhalt are missing in this database. We requested them directly from the regional statistical office of Saxony-Anhalt.
Female labor force participation (FP_{it})	Ratio of female labor force (aged 15-65) to the female resident population in the same age.	Unfortunately, data on female labor force participation is only available on the level of *Regierungsbezirke* (partly corresponding to German NUTS II regions). The data source is the microcensus (*Mikrozensus*) which provides official representative statistics of the population and the labor market in Germany. However, data is only available until 2002. We obtained the missing values by extrapolation.
Educational attainment of population ($H0_{it}$, $H1_{it}$, $H2_{it}$)	Share of employed individuals which have a job that is subject to social insurance contribution according to the level of education. Official statistics provide three levels of educational attainment: without any professional training, with a certificate of a vocational school or certificate of a university/university of applied sciences.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`). Some values for the federal country of Saxony-Anhalt are missing in this database. We requested them directly from the regional statistical office of Saxony-Anhalt.
Commuting ($COMM_{it}$)	Balance of incoming and outgoing commuters of district *i*, i.e. if the value is positive, there are more people that commute in the district than people that commute out of the district.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`). Some values for the federal country of Saxony-Anhalt and Saxony are missing in this database. We requested them directly from the regional statistical office of Saxony-Anhalt and of Saxony.
Population density ($DENS_{it}$)	Population density in individuals per km^2, values are calculated from the average population in every district divided by the area of every district.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`).
Public debts of districts ($DEBTR_{it}$)	Sum of public debts of all communities belonging to district *i* in relation to the regional domestic product of district *i*.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`).

Variable	Description	Data Source
amenities		
Number of business registrations (REG_{it})	Number of all newly registered businesses in district i.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`).
Number of overnight stays in hotels ($STAY_{it}$)	Number of people that stay over night in hotels, hostels, etc.	Regional Database of Federal Statistical Office (`www.regionalstatistik.de`).

Table B.3 Descriptions and data sources of the dependent and all explanatory variables

Appendix C

C.1 Summary statistics of matching variables

	2009Q1	2009Q2	2009Q3	2009Q4
Unemployment stock:				
min	1566	1662	1644	1597
1st quartile.	3666	3523	3379	3204
median	6380	6144	6007	5734
mean	8571	8427	8297	7864
3rd quartile	10120	10250	10130	9505
max	86040	84570	83280	79860
standard dev.	8739.74	8733.66	8647.07	8241.02
Matches:				
min	117	137.3	137	118
1st quartile	303.1	426.6	342.8	304.8
median	436.5	610.7	518.7	457.2
mean	565.1	773.6	663.5	599.1
3rd quartile	666.2	873.8	794.1	688.9
max	6504	7011	6082	6202
standard dev.	543.89	684.29	594.97	573.77
Vacancy stock:				
min	162.7	193.3	223.3	203.7
1st quartile	491.3	498.4	509.1	470
median	801.8	800.7	811,8	760.3
mean	1184	1164	1157	1118
3rd quartile	1359	1306	1312	1227
max	12120	12500	11750	11230
standard dev.	1367.47	1352.59	1305	1273.36

Table C.1 Quarterly summary statistics of unemployment stock, matches and vacancy stock for 2009

C.2 German Classification of Economic Activities (Edition 2008)

- agriculture; forestry and fishing
- industry
- industry without construction
- manufacturing
- construction
- service sector
- wholesale and retail trade; transportation and storage; accommodation and food service activities
- information and communication
- financial and insurance activities
- real estate activities
- professional, scientific and technical activities; administrative and support activities
- public administration and defense; compulsory social security; human health and social work activities
- art, entertainment and recreation; activities of households as employers; undifferentiated goods- and services-producing activities of households for own use; activities of extraterritorial organizations and bodies

C.3 Estimation results for 2008

	dependent variable: $ln(\boldsymbol{m}_t)$					
$\boldsymbol{W}_1$	contiguity	sectoral	income	GRP	social ($d^* = 25km$)	geo ($\eta = 0.02$)
$\boldsymbol{W}_2$	contiguity	sectoral	income	GRP	social ($d^* = 25km$)	geo ($\eta = 0.02$)
$ln(\boldsymbol{U}_t)$	0.75^{***}	0.72^{***}	0.72^{***}	0.72^{***}	0.72^{***}	0.75^{***}
	(64.97)	(57.51)	(58.48)	(57.79)	(57.49)	(66.64)
$ln(\boldsymbol{V}_t)$	0.02	0.04^{***}	0.04^{***}	0.04^{***}	0.04^{***}	0.01
	(1.34)	(2.87)	(2.7)	(2.77)	(2.87)	(1.18)
λ_1	-0.05^{***}	-0.04^{***}	-0.03^{***}	-0.04^{***}	-0.04^{***}	-0.05^{***}
	(-5.31)	(-3.04)	(-2.95)	(-3.02)	(-3)	(-5.02)
ρ_1	0.35^{***}	-0.35^{***}	-0.67^{***}	-0.72^{***}	-0.36^{***}	0.7^{***}
	(14.11)	(-9.67)	(-3.48)	(-5.33)	(-9.41)	(132.97)
obs.	1648	1648	1648	1648	1648	1648

Notes: t-statistics are in parentheses. λ_1 is the spatial autoregressive coefficient and ρ_1 is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table C.2 Estimation results of the $SARAR(1,1)$ matching specification with the same spatial weights for the spatial lags and spatial errors using data for 2008

	dependent variable: $ln(\boldsymbol{m}_t)$							
$\boldsymbol{W}_1$	contiguity	contiguity	contiguity	contiguity	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)
$\boldsymbol{M}_1$	sectoral	income	GRP	social ($d^* = 25km$)	sectoral	income	GRP	social ($d^* = 25km$)
$ln(\boldsymbol{U}_t)$	0.73^{***}	0.73^{***}	0.73^{***}	0.73^{***}	0.72^{***}	0.72^{***}	0.72^{***}	0.72^{***}
	(56.32)	(57.47)	(56.7)	(56.33)	(56.94)	(58.09)	(57.37)	(59.94)
$ln(\boldsymbol{V}_t)$	0.03^{***}	0.03^{**}	0.03^{**}	0.03^{***}	0.04^{***}	0.03^{***}	0.04^{***}	0.04^{***}
	(2.6)	(2.38)	(2.49)	(2.61)	(2.78)	(2.58)	(2.68)	(2.79)
λ_1	-0.04^{***}	-0.04^{***}	-0.04^{***}	-0.04^{***}	-0.03^{***}	-0.03^{***}	-0.03^{***}	-0.03^{***}
	(-3.85)	(-3.89)	(-3.9)	(-3.86)	(-2.57)	(-2.64)	(-2.69)	(-2.58)
ρ_1	-0.36^{***}	-0.76^{***}	-0.72^{***}	-0.35^{***}	-0.36^{***}	-0.75^{***}	-0.73^{***}	-0.36^{***}
	(-9.74)	(-4.19)	(-4.81)	(-10.85)	(-9.97)	(-4.07)	(-4.91)	(-11.39)
obs.	1648	1648	1648	1648	1648	1648	1648	1648

Notes: t-statistics are in parentheses. λ_1 is the spatial autoregressive coefficient and ρ_1 is the spatial autocorrelation coefficient. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table C.3 Estimation results of the $SARAR(1,1)$ matching specification using geographic weights in combination with economic distance weights (2008)

	dependent variable: $ln(\boldsymbol{m}_t)$							
$\boldsymbol{W}_1$	contiguity	contiguity	contiguity	contiguity	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)
$\boldsymbol{W}_2$	sectoral	income	GRP	social ($d^* = 25km$)	sectoral	income	GRP	social ($d^* = 25km$)
$\boldsymbol{M}_1$	contiguity	contiguity	contiguity	contiguity	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)	geo ($\eta = 0.02$)
$ln(\boldsymbol{U}_t)$	0.79***	0.75***	0.75***	0.75***	0.75***	0.75***	0.75***	0.75***
	(62.75)	(63.02)	(62.72)	(62.73)	(66.56)	(66.89)	(66.6)	(66.63)
$ln(\boldsymbol{V}_t)$	0.02	0.02*	0.02	0.02	0.01	0.02	0.01	0.02
	(1.32)	(1.74)	(1.41)	(1.46)	(1.17)	(1.4)	(1.24)	(1.34)
λ_1	−0.05***	0.001	−0.04*	−0.03	−0.06	0.15***	0.01	0.1
	(-2.22)	(-0.06)	(-1.77)	(-1.46)	(-0.56)	(2.09)	(0.06)	(1.02)
λ_2	0.0001	−0.05***	−0.01	−0.02	0.003	−0.2***	−0.06	−0.15
	(0.004)	(-2.27)	(-0.45)	(-0.69)	(0.03)	(-2.87)	(0.64)	(-1.56)
ρ_1	0.35***	0.33***	0.35***	0.34***	0.7***	0.67***	0.69***	0.68***
	(14.33)	(11.64)	(13.65)	(13.21)	(127.99)	(64.69)	(97.84)	(71.2)
obs.	1648	1648	1648	1648	1648	1648	1648	1648
Wald λ (p-value)	7.6e-07	1.5e-06	1.6e-06	4.2e-06	3.8e-06	9.9e-08	3.5e-06	3.1e-06

Notes: t-statistics are in parentheses. λ_1 and λ_2 are the spatial autoregressive coefficients associated with the spatial weights matrices W_1 and W_2, respectively. ρ_1 is the spatial autocorrelation coefficient associated with the spatial weights matrix M_1. $***$, $**$ and $*$ indicate coefficients that are significant at 1%, 5% and 10%, respectively.

Table C.4 Estimation results of the $SARAR(2,1)$ matching specification using both geographic weights and economic distance weights for the spatial lags in combination with geographic weights for the spatial error term (2008)

C.4 Choice of the optimal weighting function

Badinger and Egger (2011) propose to use a consistent estimator of $\boldsymbol{\Psi}^{-1}$ as weighting matrix $\boldsymbol{\Theta}$ to get the efficient GM estimators for ρ_1 and ρ_2. In the following, we define the different components of the estimator for $\boldsymbol{\Psi}^{-1}$. Let $\boldsymbol{\Sigma} = diag_{i=1}^{N}(\epsilon_i^2)$. Furthermore, the moment conditions in (3.19) can be written as quadratic forms in the vector $\boldsymbol{\epsilon}$, i.e.

$$M_1^{s,s'} := N^{-1}E[\boldsymbol{\epsilon}'\boldsymbol{A}_1^{s,s'}\boldsymbol{\epsilon}] \tag{C.4.1a}$$

$$M_2^{s} := N^{-1}E[\boldsymbol{\epsilon}'\boldsymbol{A}_2^{s}\boldsymbol{\epsilon}] \tag{C.4.1b}$$

where $\boldsymbol{A}_1^{s,s'} = \boldsymbol{M}_{s'}^{*}{}'\boldsymbol{M}_s^{*} - diag_{i=1}^{N}(\boldsymbol{m}_{\bullet i,s'}^{*}\boldsymbol{m}_{\bullet i,s}^{*})$ with $\boldsymbol{m}_{\bullet i,s}^{*}$ ($\boldsymbol{m}_{\bullet i,s'}^{*}$) denoting the ith column of $\boldsymbol{M}_s^{*}$ ($\boldsymbol{M}_{s'}^{*}$) and $\boldsymbol{A}_2^{s} = \boldsymbol{M}_s^{*}$. Furthermore, the $(N \times 1)$ vectors $\boldsymbol{a}_c^{s,s'}$ are defined by

$$\boldsymbol{a}_c^{s,s'} = \boldsymbol{T}\boldsymbol{b}_c^{s,s'} \tag{C.4.2}$$

where $c = 1, 2$ indicates the two groups of moment conditions $M_1^{s,s'}$ and M_2^{s} and where

$$\boldsymbol{b}_c^{s,s'} = N^{-1}E[\boldsymbol{D}'(\boldsymbol{I}_N - \sum_{m=1}^{2}\rho_m\boldsymbol{M}_m^{*}{}')[\boldsymbol{A}_c^{s,s'} + (\boldsymbol{A}_c^{s,s'})'](\boldsymbol{I}_N - \sum_{m=1}^{2}\rho_m\boldsymbol{M}_m^{*})\boldsymbol{\xi}] \tag{C.4.3}$$

with $\boldsymbol{D} = -\boldsymbol{Z}$, $\boldsymbol{I}_N$ is a N-dimensional identity matrix and $\boldsymbol{T} = \boldsymbol{F}\boldsymbol{P}$ where

$$\boldsymbol{F} = (\boldsymbol{I}_N - \sum_{m=1}^{2}\rho_m\boldsymbol{M}_m^{*}{}')^{-1}\boldsymbol{H} \tag{C.4.4}$$

and

$$\boldsymbol{P} = (N^{-1}\boldsymbol{H}'\boldsymbol{H})^{-1}(N^{-1}\boldsymbol{H}'\boldsymbol{Z})[(N^{-1}\boldsymbol{Z}'\boldsymbol{H})(N^{-1}\boldsymbol{H}'\boldsymbol{H})^{-1}(N^{-1}\boldsymbol{H}'\boldsymbol{Z})]^{-1}. \tag{C.4.5}$$

Finally, the elements of $\boldsymbol{\Psi}$ are defined as

$$\Psi_{c,c'}^{s,s',t,t'} = \frac{1}{2}N^{-1}tr([\boldsymbol{A}_c^{s,s'} + (\boldsymbol{A}_c^{s,s'})']\boldsymbol{\Sigma}[\boldsymbol{A}_c^{t,t'} + (\boldsymbol{A}_c^{t,t'})']\boldsymbol{\Sigma}) + N^{-1}\boldsymbol{b}_c^{s,s'}{}'\boldsymbol{\Sigma}\boldsymbol{b}_{c'}^{t,t'} \tag{C.4.6}$$

where $c, c' = 1, 2$; $s, t = 1, 2$ for $s = s'$ and $t = t'$, and $s = 1$, for $s' > s$. The estimate $\widetilde{\boldsymbol{\Psi}}$ is obtained through replacing the variables by consistent estimates, i.e. ρ_1, ρ_2 by $\widetilde{\rho_1}$, $\widetilde{\rho_2}$, $\boldsymbol{\Sigma}$ by $\widetilde{\boldsymbol{\Sigma}}$ and $\boldsymbol{b}_c^{s,s'}$ by $\widetilde{\boldsymbol{b}_c^{s,s'}}$.

Selbständigkeitserklärung

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Berlin, 18. Januar 2013

Franziska Lottmann